RAND

Models and Algorithms for Repair Parts Investment and Management

James S. Hodges

Arroyo Center

The research described in this report was sponsored jointly by the Navy Secretariat, Naval Air Systems Command (NAVAIR-43), and the Navy's Aviation Supply Office (ASO).

ISBN: 0-8330-1443-9

Published 1993 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
To obtain information about RAND studies or to order documents, call Distribution Services, (310) 393-0411, extension 6686

RAND

Models and Algorithms for Repair Parts Investment and Management

James S. Hodges

Prepared for the United States Navy

Arroyo Center

PREFACE

In November 1989, RAND presented a series of hypotheses and results from exploratory research to a group of senior Naval logisticians at the Naval Postgraduate School, Monterey, California, as the basis of a proposed project on Naval aviation logistics. The proposed project was to consider how lessons learned in previous RAND logistics research might apply to the Naval aviation logistics system, particularly the operation of the Naval aviation depots (NADEPs), and to consider how the detrimental effects of uncertainty on mission capability, particularly in wartime, might be offset by management adaptations. One of the most pressing problems discussed at the meeting was how to supply repair parts to the NADEPs in a timely fashion. It was decided then that one part of the project would focus on this materiel problem in an attempt to understand its dimensions and to arrive at possible solutions.

This report documents part of that research effort, describing methods for allocating investment funds among piece parts and repairable components and a method for making short-run supply decisions related to repair parts. The methods given here lend themselves to more general issues in resource allocation in repair operations, although these more general issues are discussed only briefly. The methods were developed in support of work described in MR-311-A/USN, *Using Value to Manage Repair Parts*, by Marygail K. Brauner, James S. Hodges, and Daniel A. Relles, 1993, and MR-313-A/USN, *An Approach to Understanding the Relative Value of Parts*, by Marygail K. Brauner, James S. Hodges, and Daniel A. Relles, forthcoming, and grew out of work documented in N-3473-A/USN, *Materiel Problems at a Naval Aviation Depot: A Case Study of the TF-30 Engine*, by Lionel A. Galway, 1992, and in other unpublished research. This work is part of a larger RAND project entitled *Enhancing the Logistics System: The Depot Perspective,* sponsored jointly by the Navy Secretariat, Naval Air Systems Command (AIR-43), Naval Supply System, and Aviation Supply Office.

This report should be of interest to logisticians in all Services and in commercial remanufacturing.

CONTENTS

FIGURES

TABLES

SUMMARY

The Armed Services use end-items, such as aircraft, that rely on expensive repairable components. Because of this reliance, the Services operate networks of facilities for supply and maintenance of end-items and repairable components. The Services must invest in the elements of the supply and maintenance system, allocating funds among spare repairable components, repair parts, repair capacity, second-destination transportation, and so on. At any given time, the managers of these facilities must also make operating decisions—which components to induct, which repair parts to order—that affect the performance of the maintenance and supply system and the cost of achieving that performance.

This report proposes a way to think about these investments and operating decisions, focusing on spare repairable components and repair parts. We distinguish between a long-run or investment problem and a short-run or operating problem. Our approach to both problems begins by defining the value of each part or supply action, so that the costs of either can be related to their effects, thereby permitting managers to select courses of action that maximize value given the cost of the actions. This report focuses on specific long-run and short-run problems and on the context of Service maintenance depots, but we suggest how our approach and methods can be adapted for other uses.

THE LONG-RUN (INVESTMENT) PROBLEM

The long-run or investment problem discussed here is the following.

> How should a depot make investments in repair parts—that is, how should it set authorized stock levels for repairable and consumable repair parts—to minimize the expected value of the repair pipeline for a given level of investment?

For this problem, we attribute value to units of authorized stock in terms of the effect they have on the value of the repair pipeline: efficient choices yield the lowest-value repair pipeline for a given investment in stock. Our method presumes that the depot is financially at arm's length with its suppliers. It takes as inputs

- Weapon-replaceable assembly (WRA) and shop-replaceable assembly (SRA) characteristics: depot repair induction rates, unit prices, and bills-of-material; and
- Piece part characteristics: replacement factors,[1] unit prices, and order-and-ship times (OSTs)

and computes expected awaiting-parts (AWP) times for given authorized stock levels. Once expected AWP times are available, the value of the AWP contribution to the repair pipeline is easily computed. The value of an addition to authorized stock is the amount that it decreases the value of the AWP contribution to the repair pipeline. In this method, authorized stock is viewed as an investment, a one-time buy replenished continually, that yields a return in the form of shorter AWP times and a less costly repair pipeline.

The long-run method can be applied in any kind of remanufacturing plant. Also, the method can be used as part of larger-scale considerations. such as trade-offs between different parts of the depot repair pipeline, including OSTs of repair parts, or in smaller-scale considerations, such as comparing different reliability or maintainability investments. Although

[1]Defined for each repair part, this is the chance that this part will be needed for each repair job.

this report concentrates on setting stock levels, it is likely that the greatest payoffs would come from investments in reducing OST.

THE SHORT-RUN (OPERATING) PROBLEM

We presume that in any given time period, a depot is willing to spend some budget on extraordinary actions to alleviate short-run supply problems. The short-run or operating problem, then, is the following:

> Given a repair schedule over a (short) time horizon, with given parts on the shelf and a given schedule of repair parts due in from suppliers, what actions should be taken to alter the schedule of due-ins to achieve the greatest aircraft availability for a given expenditure?

For this problem, we attribute value to supply actions—e.g., speedup of delivery of due-in items—in terms of the effect that they have on the availability of aircraft at the end of a specific time horizon. The depot selectively pays more for better supply service, targeting the expenditures to specific items that will yield the largest improvement in aircraft availability. The method takes as inputs

- WRA characteristics: bills-of-material, a schedule of WRA inductions, and the amount each unit of each WRA contributes to aircraft availability at the end of the repair horizon; and
- SRA and piece part characteristics: replacement factors, OSTs, the number of units on the shelf at the depot, and the due-in schedule

and computes the expected shortfall in WRA contributions to aircraft availability resulting from parts supply shortages. The value of a supply action is the amount that it reduces this shortfall. We presume that the short-run method would be used in conjunction with a depot induction scheduler that maximizes aircraft availability.

For both long-run and short-run problems, we give the definition of value relevant to that problem, an algorithm to maximize value for a given cost, and methods for computing value. We derive the methods in detail in Appendix A. Appendix B lists the assumptions of both methods and discusses what we know about their consequences.

ACKNOWLEDGMENTS

While developing the approach in this report, we conducted several internal seminars at RAND and received helpful suggestions from Jack Abell, John Folkeson, Don Gaver, Jean Gebman, Vince Mabert, Lou Miller, Nancy Moore, Ray Pyles, Tim Ramey, and Hy Shulman. Stephen Brady and Ken Girardini reviewed the report and made many useful comments. Jeff Stein did some computations reported in Appendix B while he was an intern at RAND in the summer of 1992. Paul Steinberg helped make this report less user-hostile. Marygail Brauner led the project in which these methods were developed; she and Dan Relles contributed ideas and comments at all stages of the work.

Over the course of RAND's work in repair parts, we have received generous help from countless civilian and uniformed Navy personnel. Most important to this report has been support from NAVAIR-43, in particular its commander, ADM Don Eaton, as well as our action officers, CDR Brian Chandler (ret.), CAPT Ryan Hansen, and CDR Bambi Smith. We have also received much appreciated help from CAPT Glenn Downer at the Navy Secretariat; CDR Kurt Hendrix, Gisela Hill, Nancy Powers, and their commanders, ADM James Eckelberger (ret.) and ADM James Davidson at the Aviation Supply Office; Tom Aton at NADEP, Jacksonville; CDR Barbara Riester (ret.), Rich Riley, and Bonnie Whitehurst at NADEP, Norfolk; and CAPT Charles Sapp, LCDR Nick Zimmon, LCDR Stan Pyle, and LCDR Dave Maddocks (ret.) at NADEP, North Island.

This work would not have been possible without the help of these people. Nonetheless, all assertions and interpretations in this report belong to the author alone.

ACRONYMS AND ABBREVIATIONS

ASO	Aviation Supply Office
AWP	awaiting parts
BCM	beyond the capability of maintenance
DLA	Defense Logistics Agency
DRIVE	Distribution and Repair In Variable Environments
E	mathematical expectation
FIFO	First In, First Out
I	daily induction rate
ICP	inventory control point
NADEP	Naval aviation depot
NAVAIR	Naval Air Systems Command
NIF	Naval Industrial Fund
NIIN	national item identification number
NIMMS	Naval Industrial Materiel Management System
NRFI	not ready for issue
NSC	Naval Supply Center
NU	no uncertainty
OST	order-and-ship time
RF	replacement factor
RFI	ready for issue
SMIC	special material identification code
SRA	shop-replaceable assembly
SU	some uncertainty
U	most uncertainty
UPA	units per application
WRA	weapon-replaceable assembly

1. INTRODUCTION

1.1. BACKGROUND

The Armed Services use end-items, such as aircraft, that rely on expensive repairable components. Because of this reliance, the Services operate networks of facilities for supply and maintenance of end-items and repairable components. The Services must invest in the elements of the supply and maintenance system, allocating funds among spare repairable components, repair parts, repair capacity, second-destination transportation, and so on. At any given time, the managers of these facilities must also make operating decisions—which components to induct, which repair parts to order—that affect the performance of the maintenance and supply system and the cost of achieving that performance. To make analytical progress on these problems, it is necessary to specify measures of merit, to measure the effects of possible investments or operating actions, and thereby to develop reasonably efficient investment and operating policies.

1.2. OBJECTIVES

This report proposes a way to proceed, focusing on spare repairable components and repair parts. We distinguish between a long-run or investment problem and a short-run or operating problem. The basis of our approach to both problems is to define the value of each part or each supply action, so that costs of either can be related to their effects and managers can select courses of action that give good value for their cost. This report focuses on specific long-run and short-run problems in the context of Service maintenance depots, but we suggest how our approach can be adapted for other uses.

1.2.1. The Long-Run (Investment) Problem

For repair parts management in Service maintenance depots, the long-run or investment problem discussed here is the following:

> How should the depot make investments in repair parts—that is, how should it set authorized stock levels for repairable and consumable repair parts—to minimize the expected value of the repair pipeline for a given level of investment?

For this problem, we attribute value to units of authorized stock in terms of their effect on the value of the repair pipeline: Efficient choices yield the lowest-cost repair pipeline for a given investment in authorized stock. The method to be described in Sections 2 and 4 takes as inputs:

- Weapon-replaceable assembly (WRA) and shop-replaceable assembly (SRA) characteristics: depot repair induction rates, unit prices, and bills-of-material; and
- Piece part characteristics: replacement factors[1] (RFs), unit prices, and order-and-ship times (OSTs)

---

[1]Defined for each repair part, this is the chance that this part will be needed for each repair job.

and computes expected awaiting-parts (AWP) times for given authorized stock levels. Once expected AWP times are available, the value of the AWP contribution to the repair pipeline is easily computed. The value of an addition to authorized stock is the amount that it decreases the value of the AWP contribution to the repair pipeline.

This long-run problem can be part of a larger supply and maintenance system trade-off: When the value of repair parts is defined in terms of repair pipeline value, investments in repair parts can be compared with other investments that reduce the value of the repair pipeline. A given amount of investment produces so much reduction of pipeline value if invested in repair parts, so much reduction if invested in second-destination transportation, and so on. An efficient investment yields the lowest-cost repair pipeline.

This definition of the value of parts can be used in other ways as well. For example, suppose the F-14's radar antenna is a problem. Which repair parts contribute most to the antenna's problem, and how can they be made less problematic? Our long-run method can be used to compute the value of each repair part on the antenna. Sorting the parts in decreasing order of value is equivalent to sorting them in decreasing order of contribution to the antenna's problem. Suppose a circuit card is the part with the highest value. If its OST is reduced, its value decreases by some amount; if the card is modified so that it breaks less frequently, its value decreases by a different amount. The relative payoffs of these or other fixes for the circuit card can be assessed by comparing their costs and the changes in value they induce.

Alternatively, this method can be used to plan capacity for warehousing and parts-handling, by sizing appropriate authorized stock and in-flows of requisitions.

Value of parts has other uses, but this suggests the possibilities. In the rest of this report, we will use value to stock parts.

Although we will focus on Service maintenance depots, the long-run method may be useful in a variety of remanufacturing plants, e.g., plants in which most major subcomponents are manufactured or remanufactured internally, or plants in which few major subcomponents are manufactured or remanufactured internally.

1.2.2. The Short-Run (Operating) Problem

We presume that in any given time period, a depot is willing to spend some budget on extraordinary actions to alleviate short-run supply problems. The short-run or operating problem, then, is the following:

> Given a repair schedule over a (short) time horizon, with given parts on the shelf and a given schedule of repair parts due in from suppliers, what actions should be taken to alter the schedule of due-ins to achieve the greatest aircraft availability for a given expenditure?

For this problem, we attribute value to supply actions—e.g., speedup of delivery of due-in items—in terms of the effect they have on the availability of aircraft at the end of a specific time horizon. The method to be discussed in Sections 3 and 5 takes as inputs:

- WRA characteristics: bills-of-material, a schedule of WRA inductions, and the amount each unit of each WRA contributes to aircraft availability at the end of the repair horizon; and
- SRA and piece part characteristics: RFs, OSTs, the number of units on the shelf at the depot, and the due-in schedule

and computes the expected shortfall in WRA contributions to aircraft availability resulting from parts supply shortages. The value of a supply action is the amount that it reduces this shortfall. As with the long-run problem, this short-run or operating problem can be posed in many different kinds of remanufacturing operations.

1.2.3. Relationship of the Long-Run and Short-Run Methods

The long-run and short-run methods discussed in later sections are not formally related: the long-run method does not take into account the kinds of adaptations that the short-run method would help to select. If the models underlying the methods were exact replicas of the actual repair and maintenance world, it is probably true that the long-run method would minimize the need for the expediting actions that the short-run method considers.

1.2.4. Attitude Toward Assumptions and Approximations

We have developed some theory and some methods based on the theory. We believe our methods should be judged by their payoff in application, not by the apparent realism of the theory, although failures of fidelity may suggest potential sources of inefficiency. Accordingly, in developing these methods we have been prejudiced toward simplicity in derivation and computation. We have paper-tested our long-run method extensively and find no systematic inefficiencies. However, these paper tests cannot address some important assumptions, and our methods' utility must be determined in field tests.

This document is intended to be self-contained, so all of the formulas given are derived in Appendix A. Many of the subsidiary results use specifications familiar from queuing theory and other areas of applied probability. We will give references to relevant work, but the exposition stands on its own.

1.3. ORGANIZATION OF THIS DOCUMENT

This report has three parts. The first part, in Sections 2 and 3, is intended for a broad logistics readership. It concentrates on aspects of the long-run and short-run methods of particular interest to executives, specifically on the analytical setup, cost issues, and the general idea of how the methods are computed. The second part, in Sections 4 and 5, is somewhat more technical. It describes the methods in a way suitable for someone who needs to implement them (e.g., someone writing computer code). The third part, the Appendices, is much more technical and is intended for readers who need to understand the methods in detail, e.g., to extend them or to apply them to other logistics problems. Appendix A derives the formulas in Sections 4 and 5. Appendix B gives the assumptions underlying the methods and discusses their consequences.

2. THE LONG-RUN (INVESTMENT) PROBLEM

The basis of our approach to the long-run problem is attributing a value to each unit of each repair part. What is an investment in the next unit of a repair part worth? Repair parts are often valued by their unit price. But unit price does not measure a part's worth: a buggy-whip has no value to the Navy, whatever it costs. We define the value of a unit of a part in terms of the benefit the Navy derives from it, where "benefit" is measured in days that end-items do not wait because of that part. (We will be more specific below.) Once value is expressed in days of end-item AWP time, it can be turned into familiar measures like "dollar value of the repair pipeline," appropriate for thinking about return on investments in parts.

Section 2.1 defines value, and Section 2.2 shows how to use value to efficiently stock parts at depots. Section 2.3 discusses how we compute value.

2.1. DEFINING THE VALUE OF PARTS

Before we can define the value of parts, we need to describe the stylized view of the depot that underlies it.

Consider a WRA, say, the radar antenna, repaired by a particular depot. The process for fixing antennas is represented in Figure 1. Units of the antenna are inducted onto an antenna repair line. Repairing them involves removing not-ready-for-issue (NRFI) repairable subcomponents, SRAs, and replacing them with ready-for-issue (RFI) SRAs, and removing and discarding NRFI piece parts and replacing them with RFI piece parts. Some types of SRAs are repaired at this depot. RFI units of those SRAs are obtained by inducting

RAND#544-1-0793

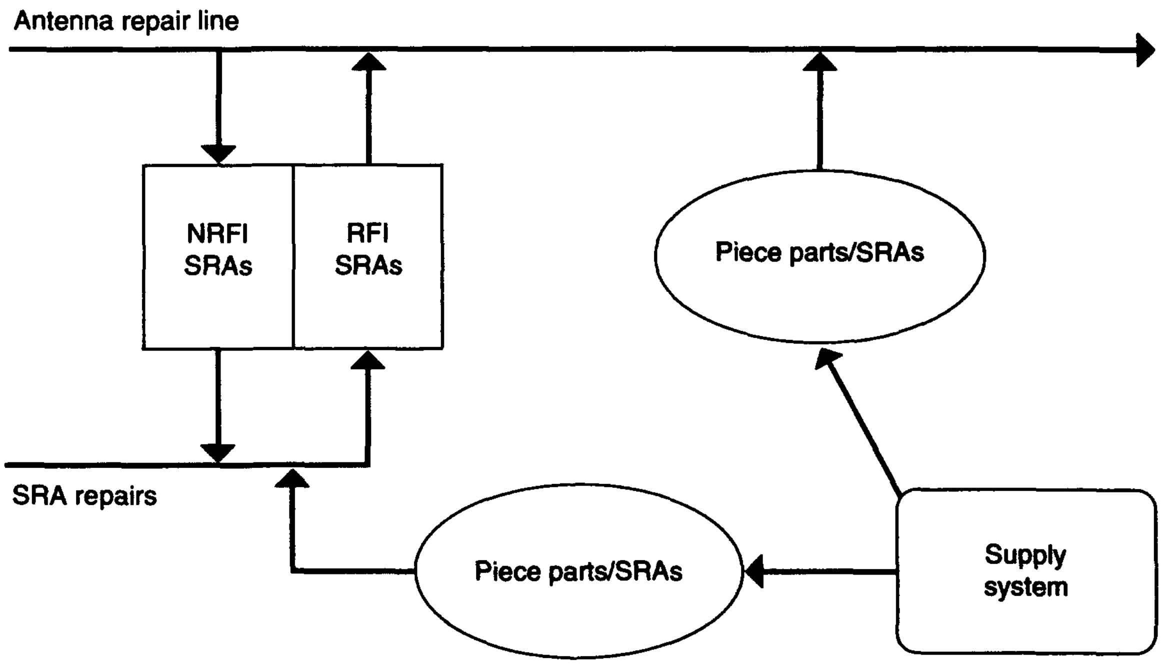

Figure 1—Schematic of Depot Repair

NRFI units onto an SRA repair line and repairing them, using piece parts and other SRAs. (Depots do not always have distinct WRA and SRA repair lines, but it is a convenient mental image.) Piece parts are either on hand at the depot or are obtained by placing requisitions to the supply system, which includes Naval Supply Centers (NSCs), supply depots of other services, Defense Logistics Agency (DLA) depots, commercial vendors, and local manufacture.

Repair parts flow out of depot stores onto antennas, and flow into depot stores in response to requisitions. The purpose of an inventory of parts—of authorized stock—is to put a buffer between the outflow onto antennas and the inflow from the supply system, so that repair jobs are inconvenienced as little as possible by requisition servicing or by repair of SRAs.

If the buffer of authorized stock is made larger or more efficient, then the AWP time of WRAs will be smaller, so fewer WRAs will be in the repair pipeline at any given time. In other words, putting more or "smarter" parts into authorized stock takes boxes out of the repair pipeline. To see how this happens, consider Figures 2 and 3, depicting a notional WRA that uses a single repair part. Figure 2 shows the situation when authorized stock of that part is zero. Each time the WRA needs a unit of the repair part, it waits an OST for it. In the figure, OST is 40 days and the repair part is needed by some unit of the WRA every 30 days. One unit of the WRA is in AWP status two-thirds of the time and two units are in AWP status one-third of the time, so the average number of units in AWP status is 4/3. Figure 3 depicts the situation when the authorized stock of the repair part is one unit. The first WRA is serviced immediately, and a part is ordered to replenish the authorized stock. Although OST is still 40 days and a unit of the WRA needs the repair part every 30 days, each WRA (except the first) waits only 10 days instead of 40. For 20 days out of every 30 no

RAND#544-2-0793

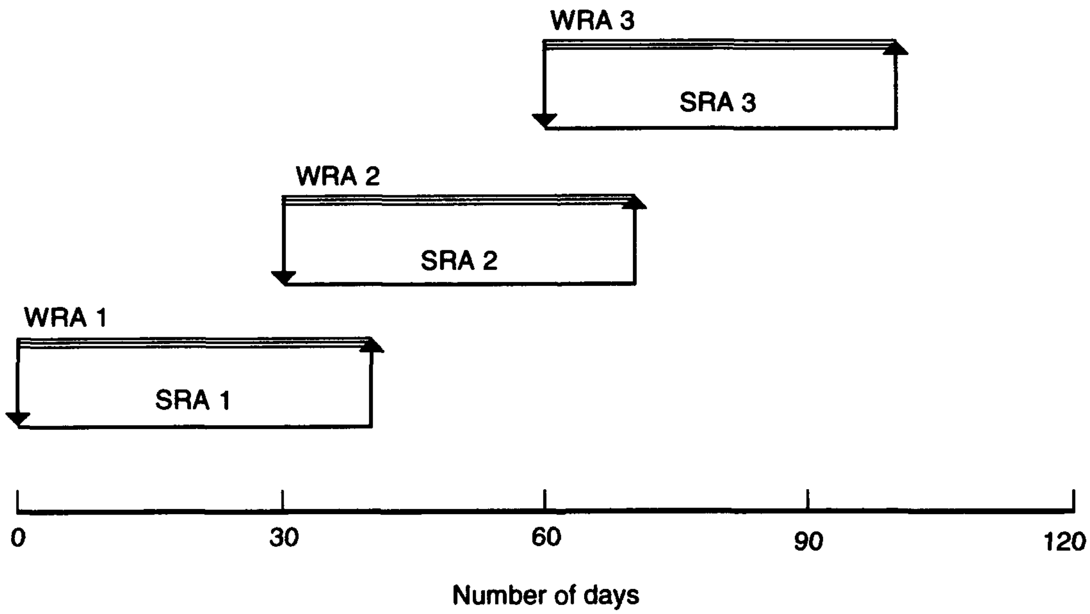

Figure 2—When Authorized Stock of the Repair Part Is Zero, Each WRA Waits an OST Each Time It Needs One

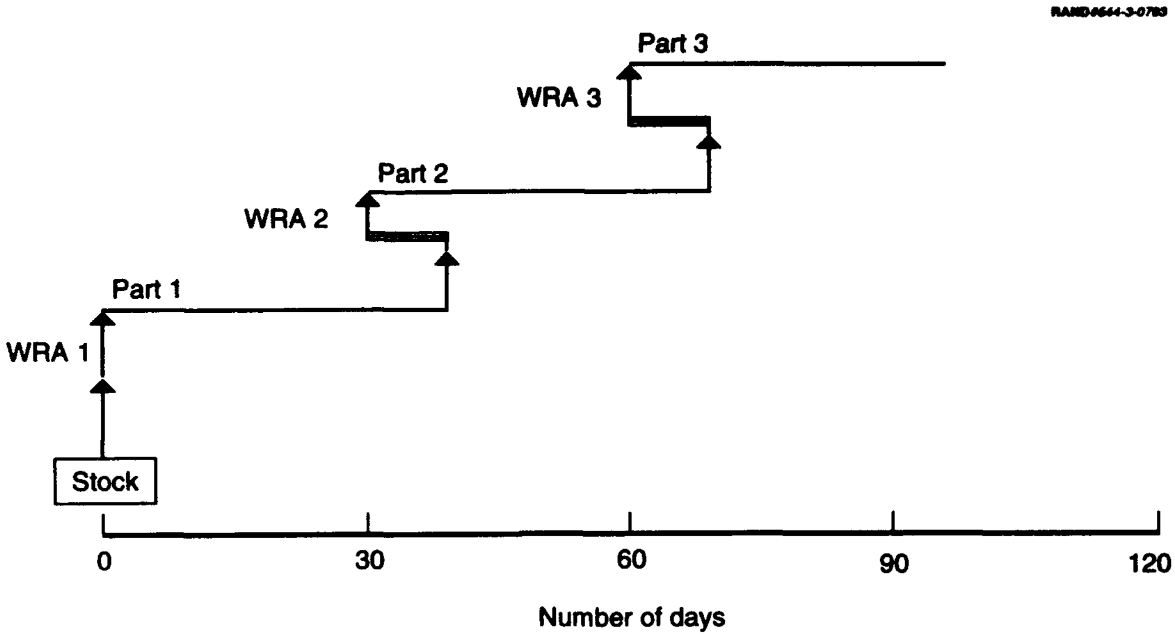

Figure 3—When Authorized Stock of the Repair Part Is One, WRAs Wait a Shorter Interval Each Time They Need One, and Fewer WRAs Are in the Repair Pipeline

units of the WRA are in AWP status, so the average number of units of the WRA in AWP status is 1/3. Stocking one unit of the repair part removes a unit of the WRA from the repair pipeline.[1]

We can now define value. A given set of stock levels yields particular expected AWP times for each WRA and a corresponding value for the repair pipeline. Add a unit of some repair part to authorized stock; the expected AWP time for the affected WRA will decrease, with a corresponding decrease in the value of the repair pipeline. The decrease in the value of the repair pipeline is the value of the unit added to authorized stock.

The foregoing discussion can be expressed mathematically. The standard formula for the value of the repair pipeline for the radar antenna is

{unit price of the antenna} × {# of antennas in the repair pipeline}

= {unit price of the antenna} ×

{antenna inductions per day} × {days of turnaround time for the antenna}.

To see this last equality, suppose an NRFI antenna is inducted each day and that two days elapse between induction of an antenna and completion of repairs. (Suppose also that two antennas can be repaired at a time.) When the depot opens its antenna repair line, it has zero antennas in work. On the first day an NRFI antenna is inducted into repair, so the depot has one antenna in work, and after the second day, it has two antennas in work. At the beginning of the third day, an RFI antenna leaves repair, but an NRFI antenna is inducted, and from then on, the depot always has two antennas in work. That is, its repair pipeline is

[1]Although 40 days may seem long, for many combinations of inventory control points (ICPs) and weapon systems (special material identification codes or SMICs, actually), the average customer wait time at depots is 40 days or greater. Reducing average OST may be a higher-payoff strategy than stocking a unit of an item.

two antennas, which is the product of the rate of antenna inductions (one per day) and the time each antenna spends in work (two days).

We assume that {days of turnaround time for the antenna} = {days of AWP time for the antenna} + {other components of turnaround time for the antenna}, and that stocks of parts affect only the AWP component of turnaround time. Among other things, this means that the processing time for an antenna is not affected by the list of parts that are needed to repair it, nor by their availability. Thus, stocks of parts affect only the contribution of AWP time to the value of the antenna repair pipeline, which is:

$$\{\text{unit price of the antenna}\} \times \{\text{antenna inductions per day}\} \\ \times \{\text{days of AWP time for the antenna}\}.$$

If this argument is applied to all WRAs, then stocks of parts affect only the contribution of AWP time to the value of the depot repair pipeline, which is the last expression summed across all WRAs. We actually compute the *expected* contribution of AWP time to the repair pipeline:

$$\text{Sum}\ (\ \{\text{unit price of the WRA}\} \times E\{\text{inductions of that WRA per day}\} \\ \times E\{\text{days of AWP time for that WRA}\}\ ),$$

where "E(X)" means "the mathematical expectation of X" and the sum is across all WRAs.

In these terms, the value of a repair part is defined as follows. For given part and antenna characteristics (to be discussed below) and given authorized stock levels, the value of the next unit of a part used to fix the antenna is the reduction in the value of the antenna's repair pipeline from adding that unit to authorized stock:

$$\{\text{unit price of the antenna}\} \times E\{\text{antenna inductions per day}\} \\ \times \Delta\ E\{\text{days of AWP time for the antenna}\},$$

where Δ E{days of AWP time for the antenna} is the change in the antenna's expected AWP time caused by the addition to authorized stock.

The value of the next unit of a part depends on the authorized levels of the other repair parts. Adding a unit of, say, a circuit card will cause different reductions in average AWP time, depending on the authorized stock levels of all repair parts, including the circuit card itself. We discuss this more explicitly in Section 2.3, below.

The notion of value discussed here implicitly assumes that the probability distribution describing part demands does not change over a long time period. In reality, the values of parts are not static: they change over time as induction rates of WRAs change, as OSTs change, and so on. Thus, the efficiency of a set of authorized stock levels also changes over time, so that an inventory that is efficient now will not be efficient in five years. This need not reflect badly on the people who buy the inventory. On the contrary, a program to improve reliability and maintainability will reduce the values of parts that are made less problematic. Depots must periodically reevaluate their authorized stock and consider changing it. This reevaluation is an important and complicated task that we do not have space to discuss here.

2.2. STOCKING PARTS TO MAXIMIZE VALUE FOR A GIVEN INVESTMENT

Given this definition of the value of the next unit of a part, it is simple to specify an algorithm to stock parts efficiently for a given investment.

Step 1. Begin with zero authorized stock for all parts.

Step 2. For each part, compute the value of the next unit of the part (we discuss how to do this in Section 2.3) and divide its value by its unit price to obtain the rate of return from stocking that unit.

Step 3. Select the part with the highest rate of return, and stock a unit of it.

Step 4. Is the total cost of authorized stock less than the total investment to be made? If so, return to step 2; if not, stop.

Steps 2 and 3 refer to evaluating and stocking the next unit of each item. It is not necessary to work with single units: Any increment to stock can be evaluated. For cheap items, such as washers, or for items with units per application (UPA) greater than one, it would usually make sense to work with increments greater than one unit. Step 4 discusses stopping the algorithm when a budget constraint is reached, but other criteria could be used to stop the algorithm. For example, stocking could continue until the average AWP time is acceptable or until the return reaches some predetermined value, such as a 2:1 return.

The stockage problem we have posed is an instance of the classic knapsack problem. Our algorithm is a heuristic solution, a so-called "greedy" algorithm, which has been extensively studied. This greedy algorithm does not, in general, maximize the objective function subject to the constraints, but it is known to produce good solutions as long as the costs of the individual increments to stock are not large relative to the budget constraint. This optimization problem could be formulated as an infinite-stage dynamic program with an average reward criterion.[2] For simplicity, we did not use this formulation.

What is the relevant measure of cost? Step 3 of the algorithm uses the unit price of each repair part as the measure of cost. To see that this is the relevant measure, consider the following. We presume that the depot is at financial arm's length with its suppliers. In other words, it must buy every part it uses, including carcasses of SRAs. If the depot had authorized stock levels of zero for all parts, it would need to buy from the supply system every part required for every repair job. If, instead, the depot has some authorized stock, it still must make the same sequence of buys from the supply system to replenish its authorized stock. The only difference is that the depot made a one-time purchase of parts up to authorized stock levels, so the repair lines receive parts more quickly. The purchase up to authorized stock levels is an outlay made in addition to the cost of parts needed for jobs, which would be incurred whatever the authorized stock levels were. The cost of authorized stock is determined by the depth of stockage and by unit prices; the return for buying that inventory is reduced AWP time. An efficient inventory is the set of authorized stock levels that makes AWP time (actually, the value of the repair pipeline) shortest (smallest) for a given investment.

Implicit in this notion of authorized stock is that it will be replenished until the parts are removed from the inventory, so the authorized stock all becomes obsolete. In reality, it will have some value as excess, so our measure of cost errs on the high side. However, if the scrap value of authorized stock is placed at some percentage of unit price, with the same per-

[2]See, for example, S.M. Ross, *Introduction to Stochastic Dynamic Programming*, New York: Academic Press, 1983, Chapter V.

centage for all items, our measure of cost produces the right ranking of additions to authorized stock. By a similar argument, we do not explicitly account for holding costs.

Our notion of costs is most sensible for weapon systems that are just being moved to organic depot repair, when the outlay for repair parts trades off directly for an outlay on WRAs to fill the repair pipeline. Our notion of costs and payoff cannot be taken literally for mature systems, because shortening the depot repair pipeline for such systems means that more RFI WRAs are available, not that fewer WRAs are procured to fill the repair pipeline. Computing actual savings for mature systems would be complicated and is beyond the scope of this report.

2.3. COMPUTING THE VALUE OF PARTS

So far we have defined value and an algorithm for stocking to maximize value without giving any details about how to compute value. This subsection discusses how we have implemented the value measure, becoming more technical as it proceeds.

Some specific assumptions needed to compute value. The ideas underlying the computation of value are illustrated by the hypothetical example in Figures 4–6. Figure 4 shows how requisitions are serviced if the depot has no authorized stock. Demands for a part, say, a circuit card for the radar antenna, occur on day 0 and on every third day afterward. If authorized stock is zero, then each time a demand for the card occurs, a unit is requisitioned from the supply system and arrives an OST later, which in Figures 4–6 is 7 days. The double line represents the amount of time each demand waits: 7 days.

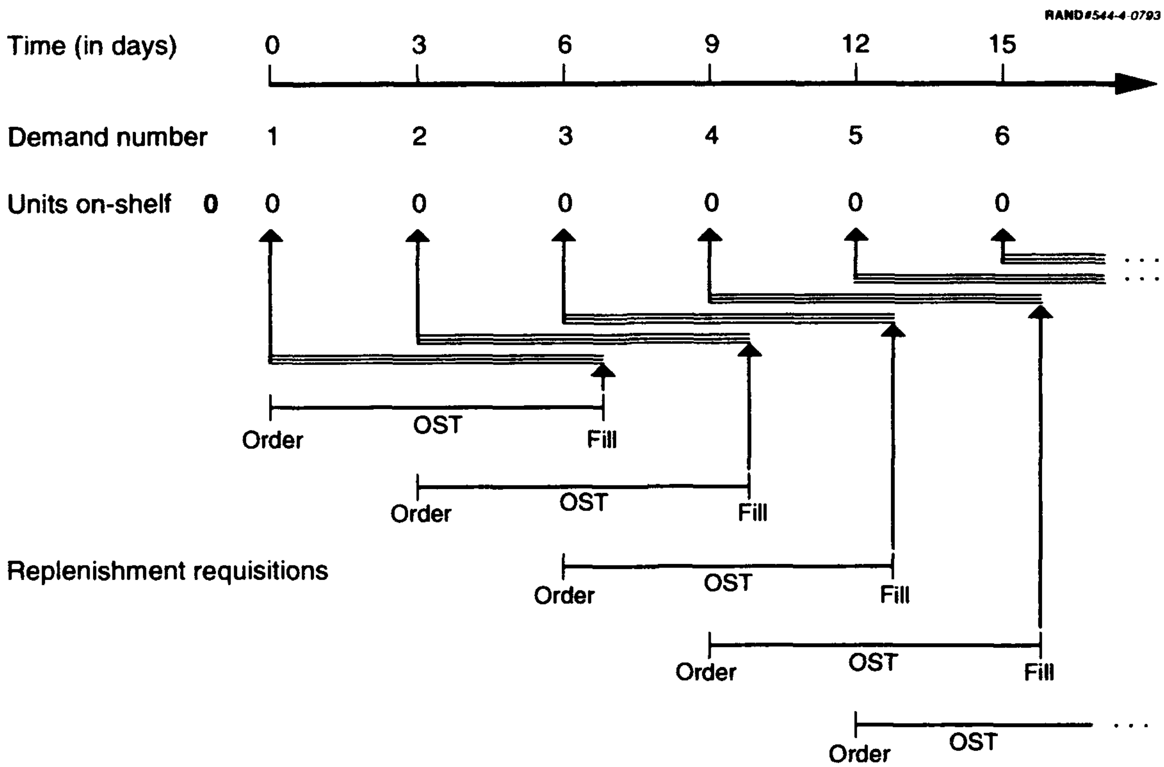

Figure 4—How Requisitions Are Serviced with No Authorized Stock

Figure 5 shows how requisitions are serviced if the depot is authorized to have one unit of stock of the card. At day 0, the depot has one unit on the shelf, which is used to fill the demand that occurs on day 0. A requisition is placed to replenish authorized stock; it is filled an OST later, on day 7. In the meantime, the second demand for the card occurs on day 3. This second demand is filled by the card that was ordered on day 0. After the second demand, a requisition is placed to replenish authorized stock; it is filled an OST later, on day 10. In the meantime, the third demand for the card occurs on day 6; and so on. As in Figure 4, the double line represents the amount of time each demand waits: with one unit of authorized stock, the wait is 4 days instead of 7.

Figure 6 shows how requisitions are serviced if the depot is authorized to have two units of stock of the card. At time 0, the depot has two units on the shelf. One of these units is used to fill the demand that occurs on day 0. A requisition is placed to replenish authorized stock; it is filled an OST later, on day 7. In the meantime, the second demand for the card occurs on day 3 and is filled by the second unit that was on the shelf at time 0. Another requisition is placed to replenish authorized stock; it is filled an OST later, on day 10. In the meantime, the third demand for the card occurs on day 6 and is filled by the unit that was ordered on day 0, so the third demand is kept waiting for a day. And so on. With two units of authorized stock, each demand waits 1 day.

This example indicates the assumptions needed to compute the value of parts. We need to describe when, in the course of an antenna repair job, the need for the circuit card becomes known, and when in the repair the card is needed. These two things are needed to define AWP time. In the example shown in Figures 4–6, we finessed these data requirements by representing the occurrence of demands for the circuit card. To use our method, we

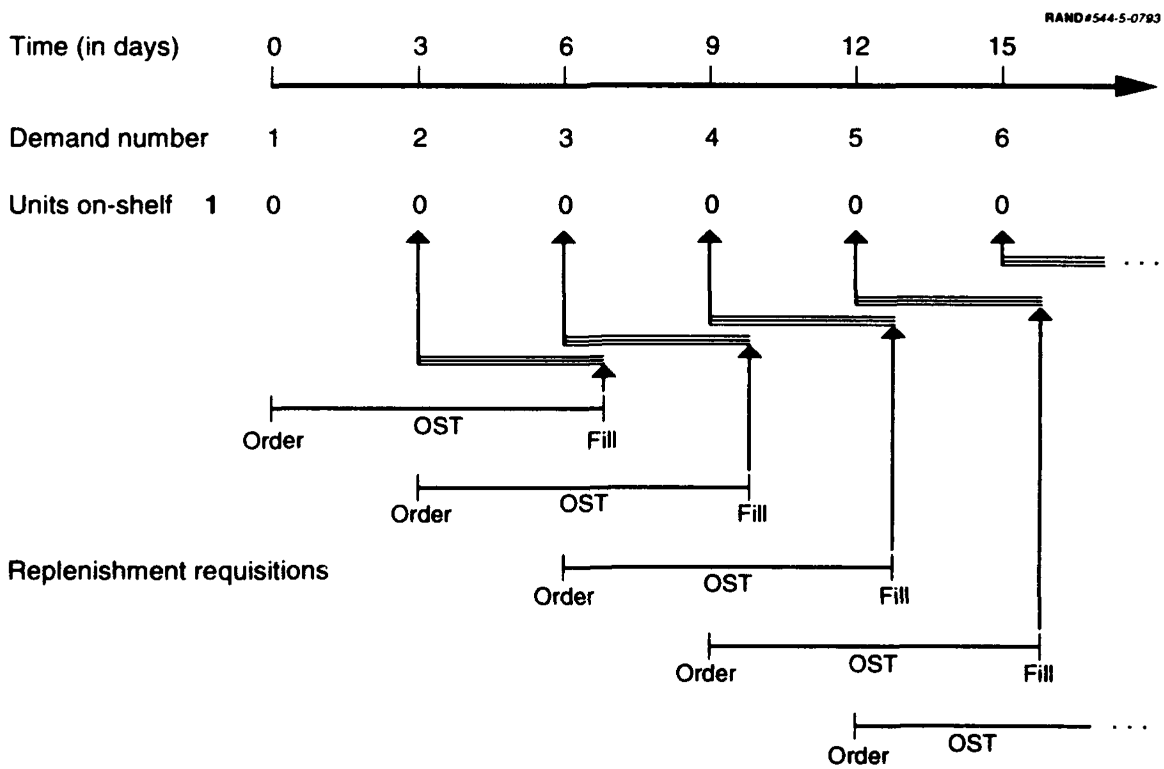

Figure 5—How Requisitions Are Serviced with One Unit of Authorized Stock

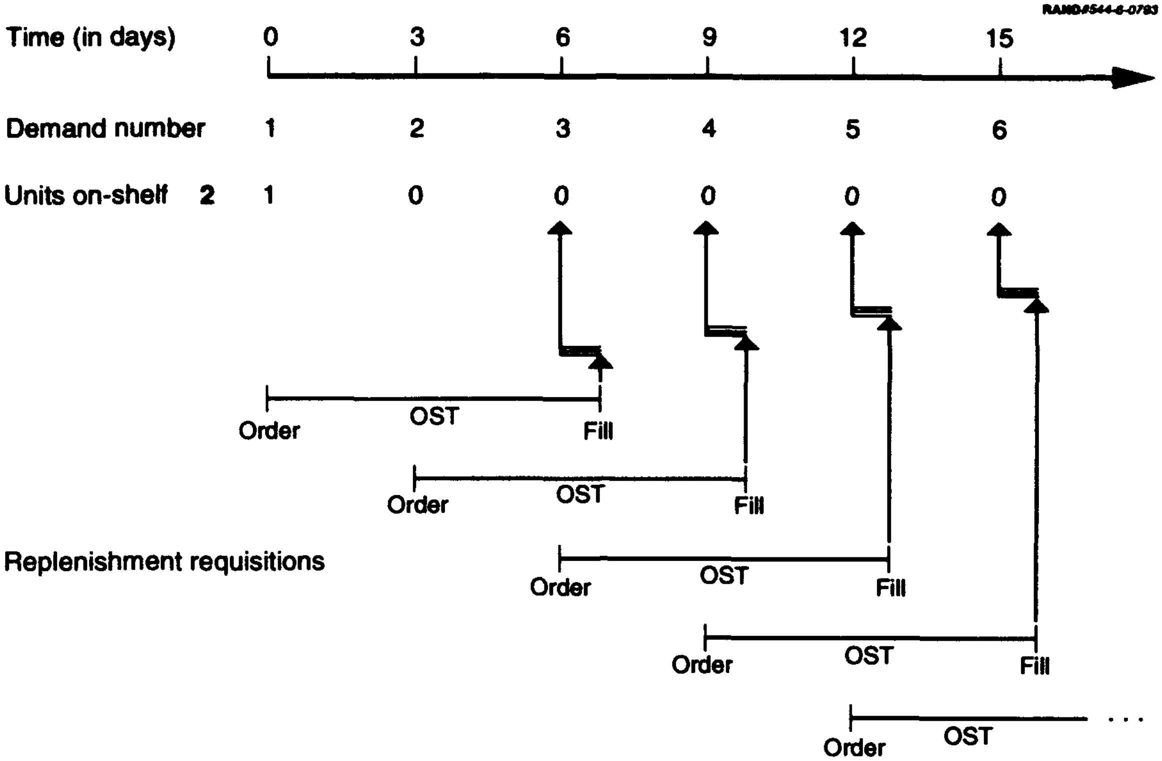

Figure 6—How Requisitions Are Serviced with Two Units of Authorized Stock

need to describe WRA inductions and how part demands arise from WRA inductions. We also need to know how depots replenish authorized stock. In Figures 4–6, each demand for the card triggered a replenishment requisition. Finally, we need to describe OST. We have used particular assumptions in each of these areas, which we discuss in the paragraphs that follow.

When part requirements for a job become known, and when in the repair those parts are needed. Computing AWP time requires assumptions about the timing of repair part demands during repair jobs. We assume that as soon as a WRA is inducted, the parts needed to repair it are known, and that they are needed immediately. These two assumptions mean that the clock measuring AWP time starts as soon as a WRA is inducted and runs until the last repair part arrives.

Neither of these assumptions is an accurate description of actual repairs, and they yield a definition of AWP time that is not consistent with Navy terminology. They will be least accurate for end-items for which processing is complicated and parts are not needed until late in the process, e.g., engine rework. Neither of our assumptions is necessary to compute value, but they greatly simplify the computation and substantially reduce its data requirements. In particular, unless the sequence and timing of actions within a job are understood and measured well, we believe it is pointless to try to replace our assumptions with assumptions having more face validity.

How depots replenish authorized stock. We assume that a replenishment requisition for part i is placed after every s^{th} demand for part i. This is often called $(S - s, S)$[3] ordering, where S is the authorized stock level. If $S = 0$—the part has no authorized stock—then s is set to 1 and an order is placed after each demand. As Figures 4–6 suggest, given

[3]See, for example, S.M. Ross, *Stochastic Processes*, New York: Wiley, 1983, p. 69.

the high OSTs currently experienced by the Navy, a depot usually will not have the authorized stock of a part on the shelf, particularly for high-demand items. Instead, units are demanded for repair jobs and replacement units are ordered; if demands occur in an orderly fashion, the bulk of authorized stock will be in the due-in pipeline. Thus, the buffer of authorized stock serves two functions. Stock on the shelf is insurance against demands for slow-moving items and against surges in demand for faster-moving items. Stock in the due-in pipeline has, in effect, been proactively ordered. If OSTs were radically shortened, most of the authorized stock would be on the shelf at any time.

Induction of WRA jobs and OST. In Section 4, we discuss different sets of assumptions about induction of WRA jobs and OST. Here is one set of assumptions (called the "U" assumptions in Section 4):[4]

- Induction of units of WRA k is a homogeneous Poisson process with rate λ_k.
- OST of part j ordered from the supply system follows a negative exponential distribution with mean $1/\delta_j$.

The other sets of assumptions incorporate less uncertainty about the timing of WRA inductions and OST. We make several other assumptions, which we discuss at the end of this section.

Computing the value of a unit of an item. Consider the circuit card going on the radar antenna. Suppose we want to assess the value of the next unit of that circuit card. For given

- Part characteristics: RFs, OSTs, unit prices, and (for SRAs) bills-of-materials
- Antenna characteristics: bill-of-materials, unit price, and daily induction rate
- Authorized stock levels for all parts,[5]

the value of the next unit of the circuit card is computed by the following steps.

1. For each SRA that is repaired locally, compute the expected AWP time for that SRA when it is repaired. This is done by:
   - For each part going on the SRA, computing the expected time the SRA waits for that part if it needs one (using formulas discussed below). This expected waiting time depends on the number of units of that part in authorized stock.
   - Plugging these and the RFs of the parts into the "tall-pole formula," which we discuss below.
2. Compute the expected AWP for the antenna. This is done by:
   - For each piece part going directly on the antenna and each SRA not repaired locally, computing the expected wait for that part if the antenna needs one (using formulas discussed below). This expected waiting time for each part depends on the number of units of that part in authorized stock.
   - For each SRA repaired locally, computing the expected wait for that SRA if the antenna needs one, using the expected AWP time computed in step 1 and

[4]For these assumptions, and assuming that in each WRA or SRA job, each part is demanded independently with constant probability, each part defines an M/M/n/∞/FIFO (First In, First Out) queue. See, for example, D. Gross and C.M. Harris, *Fundamentals of Queueing Theory*, Wiley: New York, 1974.

[5]Availability of some of these data may be problematic; the data sources we used in applying these methods for the Navy are discussed in MR-313-A/USN, *An Approach to Understanding the Relative Value of Parts*, by Marygail K. Brauner, James S. Hodges, and Daniel A. Relles.

formulas discussed below. This expected waiting time depends on the number of units of that SRA in authorized stock.

— Plugging these and the RFs of the parts that go on the WRA into the tall-pole formula.

3. Increment the circuit card's authorized stock by one unit, repeat the first two steps, and compute the change in the antenna's expected AWP time due to the extra card.
4. Compute the value of the next unit of the circuit card by multiplying the change in expected AWP time, obtained in Step 3, by the antenna's daily demand rate and unit price.

The computation just described assumes that only piece parts are used to repair SRAs. This is not true in general, and our method can be extended in an obvious way to more levels of indenture. However, as discussed in Appendix B, extending the computation to more levels of indenture may not be advisable.

The description just given assumes that we have two things: the "tall-pole formula" for computing expected AWP time given RFs and expected waits for individual parts, and a formula for the expected wait for a part if that part is needed. These are discussed in detail in Section 4 and Appendix A. The remainder of this section describes the tall-pole formula and gives some intuition for the expected wait for a part if it is needed and for how characteristics of parts and WRAs determine the value of parts.

The tall-pole formula. We use the tall-pole formula repeatedly in computing value. To understand it, consider a hypothetical WRA with four piece parts, described by Table 1.

Table 1

Part Characteristics for a Hypothetical WRA

NIIN[a]	RF	Wait If Needed
0001	0.05	40
0002	0.10	35
0003	0.50	25
0004	0.01	10

[a]National item identification number.

For each of the four parts, we have the RF and the amount of time that the artisan will wait for the part if it is needed. For now, assume that these waiting times are deterministic. If the repair job requires NIIN 0001, then no matter what other parts are demanded, the job will wait 40 days for NIIN 0001: it will be the "tall pole in the tent." If the repair job does not require NIIN 0001 but it does require NIIN 0002, then regardless of whether NIINs 0003 and 0004 are demanded, NIIN 0002 is the tall pole in the tent and the job waits 35 days. If neither NIIN 0001 nor NIIN 0002 is required, but NIIN 0003 is, then it is the tall pole and the job waits 25 days. If only NIIN 0004 is demanded, the job waits 10 days. Therefore, the mathematical expectation of the time the job waits for parts is

$$
\begin{aligned}
\text{E(wait)} &= \text{Sum(\{Wait if part i is the tall pole\}} \\
&\quad \times \text{Probability\{part i is the tall pole\})} \\
&= 40 \times \text{Pr(Need NIIN 0001)} \\
&\quad + 35 \times \text{Pr(Need NIIN 0002, don't need 0001)} \\
&\quad + 25 \times \text{Pr(Need NIIN 0003, don't need 0001 or 0002)} \\
&\quad + 10 \times \text{Pr(Need NIIN 0004, don't need 0001, 0002, or 0003)} \\
&\quad + 0 \times \text{Pr(don't need any parts)} \\
&= 40 \times 0.05 + 35 \times 0.10 \times 0.95 + 25 \times .50 \times .95 \times .90 \\
&\quad + 10 \times .01 \times .95 \times .90 \times .50 \\
&= 16.06
\end{aligned}
$$

This is the tall-pole formula for the expected amount of time a WRA (or SRA) waits for repair parts. If the waiting times for individual parts are deterministic, the tall-pole formula gives the exact expected value. If the waiting times for individual parts are stochastic—so that the column "wait if needed" in Table 1 gives the *expected* wait if needed—then the stochastic process yielding the WRA's wait for parts is much more complicated and the tall-pole formula is approximate. The complication arises because first, the parts required for the repair are determined stochastically, then waiting times for the needed parts are determined stochastically, then the WRA's waiting time is determined as the maximum of those waiting times. As the number of parts on the WRA grows, it becomes much more time-consuming to compute the expectation of this maximum. The tall-pole formula approximates the exact computation. The quality and consequences of this approximation are discussed in Appendix B.

How WRA and part characteristics determine the value of parts. Consider a hypothetical radar antenna composed of two parts, a circuit card and a bolt. Recall the definition of the value of the next unit of the circuit card:

$$
\begin{gathered}
\text{\{unit price of the antenna\}} \times \text{\{antenna inductions per day\}} \\
\times \Delta\ \text{E\{days of AWP time for the antenna\},}
\end{gathered}
$$

where Δ E(days of AWP time for the antenna) is the reduction in the antenna's expected AWP from adding a unit of the circuit card to authorized stock. The card's value depends on:

- The unit price of the antenna;
- The induction rate of the antenna; and
- How the characteristics of the circuit card and the bolt affect Δ E{days of AWP time for the antenna}.

Value is proportional to the unit price of the antenna. Suppose the circuit card and another part, say, a diode, could each knock one day off the AWP time of, respectively, the antenna and a converter-programmer. If the antenna and converter-programmer fail at the same rate, then the Navy benefits more by saving the day on the more expensive WRA.

The induction rate of the antenna and the characteristics of the circuit card and the bolt have more complicated effects on the value of the next unit of the card. To understand these effects, consider the tall-pole formula for our hypothetical radar antenna. If the expected wait for the bolt is longer, the tall-pole formula yields

$$\text{E(antenna AWP)} = \text{RF(bolt)} \times \text{E(wait for bolt if needed)} + (1 - \text{RF(bolt)}) \times \text{RF(card)} \times \text{E(wait for card if needed)}.$$

If the expected wait for the card is longer, the tall-pole formula yields

$$\text{E(antenna AWP)} = \text{RF(card)} \times \text{E(wait for card if needed)} + (1 - \text{RF(card)}) \times \text{RF(bolt)} \times \text{E(wait for bolt if needed)}.$$

Thus, Δ E(antenna AWP) =

- RF(card) $\times$ Δ E(wait for card if needed), if the wait for the card is longer, and
- (1 – RF(bolt)) $\times$ RF(card) $\times$ Δ E(wait for card if needed), if the wait for the bolt is longer.

These two formulas show how the value of the card depends on the bolt's characteristics. If the expected wait for the card (given its authorized stock level) is longer than the expected wait for the bolt (given its authorized stock level), the value of the next unit of the card does not depend explicitly on the bolt's characteristics. If the expected wait for the bolt is longer than the expected wait for the card, the value of the card is determined by the bolt's RF. For example, if the expected wait for the bolt is 30 days and the expected wait for the card is 25 days, *and* antennas always need the bolt—that is, RF(bolt) = 1—then the next unit of the card has zero value, because the antenna waits longer for the bolt than for the card regardless of whether another card is stocked. If only half the antenna jobs need the bolt—RF(bolt) = 0.5—then the AWP times of the other half of the antenna jobs can potentially be reduced by reducing the waiting time for the circuit card, so the next unit of the circuit card has some value.

The effects of the induction rate of the antenna and the circuit card's characteristics are somewhat more complicated. From the last formula given, the value of the next card is proportional to:

$$\{\text{Induction rate for antenna}\} \times \text{RF(card)} \times \Delta\, \text{E(wait for card if needed)}.$$

(Actually, this is true only as long as the expected wait for the card and the expected wait for the bolt remain in the same order; when the stock of the card increases to the point at which its expected wait becomes smaller than the bolt's expected wait, the card's value changes discontinuously.) The third factor in this expression, Δ E(wait for card if needed), is determined by the assumptions made about antenna inductions and OST. For the U assumptions given in Section 4,

$$\Delta\text{E(wait for card if authorized stock = n)} = \left(\frac{\text{RF(card)} \times \text{I(antenna)}}{\text{RF(card)} \times \text{I(antenna)} + 1/\text{E\{OST(card)\}}}\right)^{n},$$

where I(antenna) is the daily induction rate of the antenna and E{OST(card)} is the expected OST of the card. Thus

$$\Delta E(\text{wait for card if needed}) =$$

$$E(\text{wait for card if authorized stock} = n) - E(\text{wait for card if authorized stock} = n + 1)$$

$$= \frac{(RF(card) \times I(antenna))^n}{(RF(card) \times I(antenna) + 1/E\{OST(card)\})^{n+1}},$$

so the value of the next card is proportional to:

$$\left(\frac{RF(card) \times I(antenna)}{RF(card) \times I(antenna) + 1/E\{OST(card)\}}\right)^{n+1}.$$

As the authorized stock of the card increases, the value of the next unit decreases—that is, the marginal return to stocking the card is decreasing. As the expected OST of the card increases, the value of the next unit of the card increases. The derivative of value with respect to RF is positive, so the card's value increases with its RF. (The second derivative is negative, so that the increase in value with RF slows as the RF becomes larger). Symmetrically, as the antenna's induction rate increases, so does its value. (The second derivative of value with respect to induction rate is also negative.) If the antenna never breaks—I(antenna) = 0—or if the antenna breaks but the circuit card never does—RF(card) = 0—there can be no value to stocking the circuit card. If either the antenna's induction rate or the card's RF is increased, an antenna job will wait longer each time a circuit card is demanded, so the value of the next unit of the card is increased.

Miscellaneous other assumptions. Our method uses other assumptions that may be important. We assume the following.

- The depot has adequate capacity to repair the WRAs that are scheduled.
- The number of units of a WRA in work is small relative to the total number of units owned by the Navy. The effect of this assumption is that the induction rate of WRAs is not affected by the number in the repair pipeline.
- Due-in parts are not used opportunistically, that is, we assume FIFO. In Figure 4, this means the part ordered for demand 1 is used to fill demand 1, even if using it to fill demand 2 would make a WRA available sooner.
- SRA repair time = process time + AWP time, where process time is unaffected by the availability of parts.
- No cannibalization.
- No condemnation of WRAs or SRAs.
- UPA is 1.
- No common items.

A complete list of assumptions of the long-run methods is given in Appendix B. In Section 4, we consider three specific probabilistic models for WRA inductions, OST, and demands for parts for individual WRA or SRA jobs.

As noted above, each individual part follows a queueing process, with the specific process depending on the specific probabilistic forms for WRA inductions, OST, and the occurrence of part demands for individual WRA or SRA jobs. The new wrinkle in our setup is that for each WRA or SRA repair job, the AWP time is the maximum across several queues selected stochastically from a larger group of queues, corresponding to the parts of which the WRA or SRA is composed.

3. THE SHORT-RUN (OPERATIONAL) PROBLEM

As with our long-run method, the basis of our short-run method is the attribution of value. For the long-run problem, we attribute value to increments to authorized stock in terms of the amount they reduce the value of the repair pipeline. For the short-run problem, we attribute value to supply actions affecting repair parts in terms of their contribution to aircraft availability over a short time horizon. (An example of a supply action is speeding up a due-in.) Section 3.1 defines the value of a supply action and Section 3.2 shows how to use value to efficiently select supply actions. Section 3.3 discusses how to compute value for the short-run problem. This section parallels Section 2 to emphasize the common ideas.

3.1. DEFINING THE VALUE OF SUPPLY ACTIONS

The stylized view of depot operations underlying our short-run method is the same as in Section 2.1 and Figure 1. In Section 2.1, however, depot operations were timeless in the sense that all the related phenomena were assumed to be stable over time. In this section we focus on a specific time period—say, a two-week period—and on how repair parts affect aircraft availability at the end of that time period. To do this, we must add some details to the view of depot operations given in Section 2.1.

Consider a specific WRA, the radar antenna, and the depot at which it is repaired. We assume that for some time period, say, two weeks, the depot has determined how many antennas it will induct and when it will induct them.[1] We also assume that enough NRFI antennas will be available to execute this schedule.[2] An example of this would be that the depot plans to induct six NRFI antennas, on days 0, 2, 4, 6, 8, and 10. Further, we assume that for each repair part used to fix the antenna, depot materiel managers know how many units are on the shelf, how many units are due in from the supply system, and when they are due. For parts ordered under an (S – 1, S) scheme, the number of units on the shelf plus the number of units that are due in equals the authorized stock level. Under this setup, all supply actions can be expressed as changes to due-dates of units that are on the schedule or as additions of units to the schedule. Our short-run method is concerned with evaluating such changes to the due-in schedule.

We assume that the depot has a fixed number of antennas scheduled for induction on specific dates over a fixed future period. Let that number of antennas be R. We assume further that if the r^{th} antenna on the schedule is repaired in a timely fashion, it makes a particular contribution, c_r, to aircraft availability, and that the total contribution to availability made by the antennas on the schedule is the sum of their individual contributions.[3] The antennas and their contributions to availability are depicted in Figure 7. If the r^{th} antenna cannot be repaired in a timely way (which we will define shortly), then it does not make a

[1]The length of this time period is determined, to some extent, by OST. If OSTs are all two days, it makes little sense to consider a two-week horizon.

[2]Our short-run methods were developed to be compatible with a depot induction scheduler called DRIVE (Distribution and Repair In Variable Environments), described in *DRIVE (Distribution and Repair In Variable Environments): Enhancing the Responsiveness of Depot Repair*, RAND, R-3888-AF, 1992, by J.B. Abell, L.W. Miller, C.E. Neumann, and J.E. Payne, and *DRIVE (Distribution and Repair In Variable Environments): Design and Operation of the Ogden Prototype*, RAND, R-4158-AF, 1992, by L.W. Miller and J.B. Abell.

[3]In DRIVE, an antenna's contribution to availability is the difference between log(probability of meeting aircraft availability goals with this antenna on the schedule) and log(probability of meeting aircraft availability goals without this antenna on the schedule), where the aircraft availability goals are specified by operating organizations. These c_i are determined by aircraft operating tempos, WRA break rates, and beyond-the-capability-of-maintenance (BCM) rates.

RAND#544-7-0793

Sequence of antennas	1	2	3	. . .	R – 2	R –1	R
Contribution to availability	C_1	C_2	C_3	. . .	C_{R-2}	C_{R-1}	C_R

Total contribution to availability = $C_1 + C_2 + C_3 + \ldots + C_{R-2} + C_{R-1} + C_R$

Figure 7—The Antenna Repair Schedule and the Antennas' Contribution to Aircraft Availability

contribution to aircraft availability. This is depicted in Figure 8: The last two antennas on the schedule cannot be repaired in time, and so the total contribution to availability made by the antennas on the schedule is reduced by $c_{R-1} + c_R$.

We assume that there is a time t such that if an antenna is in AWP status at time t, it cannot be fixed soon enough to contribute to aircraft availability. This is depicted in Figure 9: If the necessary parts are not available at time t, the antenna cannot be repaired, shipped forward, and installed on the aircraft by the operating organization's drop-dead date. Supply actions yield value by reducing the number of antennas that are in AWP status at time t.

We can now define the value of a supply action. A given state of part supply—that is, given numbers of parts on the shelf and a given due-in schedule—yields a particular expected number of antennas in AWP status at time t. A supply action, such as speeding up a due-in, decreases the expected number of antennas in AWP status at time t. The antennas that are no longer in AWP status contribute to aircraft availability. The additional contribution to availability is the value of the supply action.

This can be expressed mathematically. Suppose that for a given state of parts supply, the expected number of antennas in AWP status at time t is α. Because α will usually not be an integer, let $\alpha = A + d$, where A is an integer and $d = \alpha - A$ is between 0 and 1. Suppose

RAND#544-8-0793

Sequence of antennas	1	2	3	. . .	R – 2	$\cancel{R-1}$	$\cancel{R}$
Contribution to availability	C_1	C_2	C_3	. . .	C_{R-2}	$\cancel{C_{R-1}}$	$\cancel{C_R}$

Now, total contribution to availability = $C_1 + C_2 + C_3 + \ldots + C_{R-2}$

Figure 8—If Some Antennas Cannot Be Fixed in Time, They Do Not Contribute to Availability

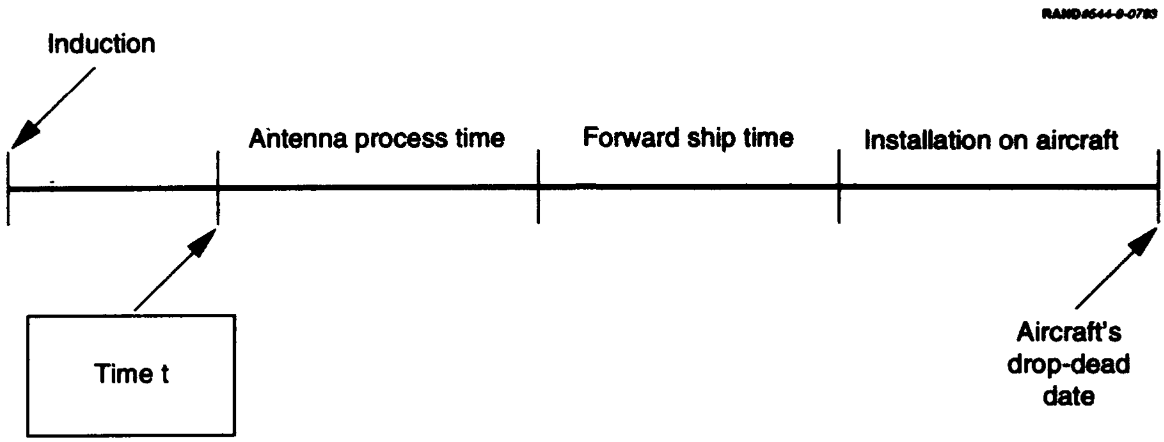

Figure 9—Assume There Is a Time t Such That If an Antenna Is in AWP Status at Time t, It Cannot Be Fixed Soon Enough to Contribute to Aircraft Availability

that a supply action reduces the expected number of antennas in AWP at time t to β, and by analogy, let $\beta = B + e$, where B is an integer and $e = \beta - B$ is between 0 and 1. Then the value of the supply action is

$$\sum_{r=1}^{R-B-1} c_r + (1-e)\, c_{R-B} - \left\{ \sum_{r=1}^{R-A-1} c_r + (1-d)c_{R-A} \right\},$$

which is the increase in the antenna's contribution to availability due to the supply action.

This definition of value depends on the given state of parts supply. Adding a unit of, say, a circuit card to the due-in schedule, arriving on day 10, will cause different increases in aircraft availability depending on the state of supply of other parts and of the circuit card itself. We will discuss this more explicitly in Section 3.3. Also, this definition of value is a short-run notion. The values of supply actions change over time as the state of parts supply changes and as the c_i are changed by variation in aircraft operating tempos, WRA break rates, and BCM rates. Thus, the efficiency of a given set of supply actions also changes over time, so that a set of supply actions that is efficient one week may not be efficient a week later.

It is possible to define value for the short run as a weighted average of changes in AWP time, in a direct analogy to the long-run measure of value. We will not describe this alternative definition of value further.

3.2. SELECTING SUPPLY ACTIONS TO MAXIMIZE AIRCRAFT AVAILABILITY FOR A GIVEN COST

Given this definition of the value of a supply action, it is straightforward to specify an algorithm for selecting supply actions to maximize aircraft availability.

Step 1. Begin with a particular state of supply and no supply actions.
Step 2. For each supply action under consideration, compute its value (we discuss how to do this below) and divide by its cost to obtain its rate of return.
Step 3. Select the supply action with the highest rate of return.

Step 4. Is the total cost of supply actions less than the amount you are willing to spend? If so, return to step 2; if not, stop.

Step 2 presumes the existence of a collection of possible supply actions, each of which is to be evaluated. The possible supply actions should involve single parts, or at most a collection of parts that all go on a single WRA. Apart from this, we have little general guidance to offer. Step 4 describes stopping the algorithm when a budget constraint is reached, but other criteria could be used to stop the algorithm, for example, a cap on the total number of supply actions or a minimum aircraft availability. The short-run problem is, like the long-run problem, a particular instance of the classic knapsack problem, and the algorithm we propose is a "greedy" algorithm. The short-run problem could be formulated as a finite-stage dynamic program,[4] but for simplicity, we did not do so.

What is the relevant measure of cost? Step 3 of the algorithm uses the cost of the supply action. By "cost of the supply action," we mean the additional cost incurred by resorting to the supply action instead of allowing ordinary procedures to take their course. If the depot takes no supply actions, it pays for parts obtained by ordinary means, for example, shipped under the usual delivery-time standards. If the depot has a part shipped by an express shipper instead of by the usual carrier, the cost attributable to that supply action is the *difference* between the cost of express shipping and the cost of the usual carrier. This difference is the cost to be used in Step 3.

3.3. COMPUTING THE VALUE OF SUPPLY ACTIONS

We have defined value and an algorithm for selecting supply actions to maximize value. This section discusses how we have implemented the value measure, becoming more technical as it proceeds.

Some specific assumptions needed to compute value. We use the same basic ideas to compute value for the short-run problem as we did for the long-run problem. Thus, we need assumptions similar to those discussed in Section 2, in particular, regarding:

- When part requirements for a job become known, and when in the repair those parts are needed;
- How depots replenish authorized stock; and
- Induction of WRA jobs and OST.

For the first two bullets, we use the same assumptions as were used for the long-run method. We assume that, as soon as a WRA is inducted, the parts needed to repair it are known, that they are needed immediately, and that replenishment follows an (S – s, S) rule. This replenishment policy has an important implication in the short-run problem: without specific supply actions by depot personnel, parts that are used for repair will be replenished automatically by the (S – s, S) rule. Thus, if three units of a circuit card are on the shelf or due in, but the current schedule of antenna repairs requires five units, the last two units will be requisitioned from the supply system through the (S – s, S) rule unless a supply action is made to order them more promptly.

We assume that the number of WRAs to be inducted during the planning horizon is specified—previously, we called it R—and that the dates they will be inducted are also specified. We assume that for each repair part, the depot knows the number of units on the shelf at time 0, and the number of units due in and their due dates. The formulas given in Section

[4] See, for example, Ross, *Introduction to Stochastic Dynamic Programming*, op. cit.

5 and in Appendix A assume that the due dates and OST are deterministic. This is not necessary, and the derivations will indicate how probability distributions can be introduced.

We make several other assumptions; these are discussed at the end of this section.

Computing the value of a supply action. This computation has fewer distinct steps than the computation for the long-run methods. For given

- Part characteristics: RFs, OSTs, and (for SRAs) bills-of-material;
- WRA characteristics, summarized in the c_r; and
- Available supply actions, represented by their effect on the due-in schedule and by the cost increment over ordinary supply,

the value of, say, speeding up a due-in of a circuit card on the radar antenna is computed by the following steps:

1. For the current state of supply, compute the expected contribution of radar antennas to aircraft availability.
2. Speed up the due-in of the circuit card and recompute the expected contribution of radar antennas to aircraft availability.
3. The change in expected contribution to availability is the value of speeding up the due-in.

This description assumes that we can compute the expected number of antennas that will be in AWP status too late to contribute to aircraft availability. This is discussed in detail in Section 5 and in Appendix A. The remainder of this section describes the computation in less detail, to convey some intuition for how it works.

Computing the change in the expected number of antennas that will be in AWP status too late to contribute to aircraft availability. Recall that the value of a supply action for antenna parts (from Section 3.1) is

$$\sum_{r=1}^{R-B-1} c_r + (1-e)\,c_{R-B} - \left\{ \sum_{r=1}^{R-A-1} c_r + (1-d)c_{R-A} \right\}.$$

This can be approximated as

$$c^{(m)} \times \Delta\, E\{\text{number of antennas in AWP status at time } t\}$$

where $c^{(m)}$ is the average of the contributions of the marginal antennas, $c_{R-A}, \ldots, c_{R-B}$, and Δ E{number of antennas in AWP status at time t} is the change in the expected number of antennas that cannot be repaired in time to contribute to aircraft availability.

Consider the value of a supply action affecting a circuit card on the antenna. Characteristics of the antenna contribute to the value of the supply action through $c^{(m)}$. Because the specific contribution of antenna characteristics will depend on how the c_r are defined, we can speak only in generalities. In general, as the antenna's systemwide stock position becomes poorer, $c^{(m)}$ should become larger, so the value of a given supply action will also become larger. Similarly, $c^{(m)}$ should increase with the break rate of the antenna.

Characteristics of the circuit card and the other parts on the antenna contribute to the value of a circuit card supply action through Δ E{number of antennas in AWP status at time t}. Suppose the antenna has only two parts on it, the card and a bolt. Section 5 shows that

E{number of antennas in AWP status at time t}

$$= R - \sum_{r=1}^{R} \Pr \text{ (card doesn't keep the } r^{th} \text{ antenna waiting past time t)}$$
$$\times \Pr \text{ (bolt doesn't keep the } r^{th} \text{ antenna waiting past time t)}.$$

Therefore, the change caused by a supply action affecting the card is

Δ E{number of antennas in AWP status at time t}

$$= \sum_{r=1}^{R} \Pr \text{ (bolt doesn't keep the } r^{th} \text{ antenna waiting past time t)}$$
$$\times \Delta \Pr \text{ (card doesn't keep the } r^{th} \text{ antenna waiting past time t)}.$$

where Δ Pr(card doesn't keep the r^{th} antenna waiting past time t) is the change caused by the supply action affecting the circuit card. The effect of the bolt is simple: if the bolt's supply position is worsened, so that each antenna on the schedule is more likely to wait for a bolt past time t, then a given Δ Pr(card doesn't keep the r^{th} antenna waiting past time t) has a smaller effect on availability and therefore the supply action has lower value. The bolt's supply position is worsened if either its RF or its OST is increased.

Characteristics of the circuit card have their effect through Δ Pr(card doesn't keep the r^{th} antenna waiting past time t). These effects are difficult to explain without deriving this change in probability, so we defer interested readers to Appendix A. However, value increases with the RF of the card.

Miscellaneous other assumptions. The short-run method uses many of the same assumptions that were used for the long-run method. In particular, we assume

- SRA repair time = process time + AWP time, and only the latter can be affected by supply actions.
- The depot has adequate capacity to repair the WRAs that are scheduled.
- Due-in parts are not used opportunistically, that is, we assume FIFO.
- No cannibalization.
- No condemnation of WRAs or SRAs.
- UPA is 1.
- No common items.

The short-run methods require some additional assumptions. These include

- The depot has a menu of supply actions.
- For each supply action on the menu, the depot knows what the resulting due date will be, and how much the supply action will cost.

These last two assumptions depend on the visibility of previously placed requisitions and on costing conventions that may be beyond the depot's control. A complete list of assumptions of the short-run methods is given in Appendix B.

As for the long-run problem, each individual part follows a queueing process. In the long-run problem, the long-run properties of the queues are of interest, but in the short-run problem, the short-run properties are relevant.

4. FORMULAS FOR EXPECTED AWP TIME IN THE LONG-RUN PROBLEM

Section 2 described our method for stocking parts according to value, omitting the formulas for expected AWP times. Expected AWP time is determined by assumptions made about arrival of WRA jobs, OST, and so on, as discussed in Section 2. In this section, we give formulas for three sets of such assumptions. Section 4.1 discusses the rationale for having three sets of assumptions, and Sections 4.2, 4.3, and 4.4 give the formulas. Appendix A derives the formulas, and Appendix B lists all of the assumptions underlying these formulas.

4.1. WHY THREE SETS OF ASSUMPTIONS?

We designate the three sets of assumptions by

- *No Uncertainty* (NU): no uncertainty in anything;
- *Some Uncertainty* (SU): the only uncertainty is about which repair parts will be needed for each job;
- *Most Uncertainty* (U): uncertainty about job inductions, OST, and which repair parts will be needed for each job.

Why not always use U? U includes assumptions about OST and times between WRA jobs that may introduce inefficiency if jobs arrive regularly and OST is reliable. Assumptions intermediate between SU and U can be specified, for example, including uncertainty about OST but not about inductions. Appendix A suggests how to derive the needed formulas. Also, NU may be simpler to use than U, so even if U is more realistic, as it often will be, it may not be more cost-effective. SU is more complicated than either of the other two. It is included partly because some users may find it appropriate, but mostly for completeness.

The next three subsections give formulas for expected AWP times for NU, SU, and U. The specifications for the three sets of assumptions will be displayed in Tables 2–4 so that they can be compared readily. In all cases, we give formulas for (S –1, S) ordering and (S – s, S) ordering. We take s as given and do not optimize over it.

4.2. FORMULAS FOR EXPECTED WAITING TIMES UNDER NU ASSUMPTIONS

Recall the setup discussed in Section 2: WRA jobs are inducted for repair; they call for SRAs and for piece parts that go directly on the WRA; SRAs are repaired, and their repair consumes piece parts; and piece parts come from in-house stocks or from the supply system. The NU assumptions about these processes are particularly simple:

Outflow of parts

- Inductions of WRA k occur equally spaced in time at a known rate of D_k per day, so the time between inductions is $1/D_k$ days.
- Demands for SRAs occur at equal known intervals determined by their RFs; if p_i is SRA i's RF, then a demand for it occurs regularly, once every $1/(p_iD_k)$ days.
- Demands for piece parts going directly on the WRA occur at equal known intervals determined by their RFs; if p_j is piece part j's RF, a demand for it occurs regularly, once every $1/(p_jD_k)$ days.

- **Demands for piece parts used to fix SRAs occur at equal known intervals determined by their RFs and the SRA's RF; if p_i is the SRA's RF and q_j is the RF of piece part j used to fix it, a demand for piece part j occurs regularly, once every $1/(p_i q_j D_k)$ days.**

Inflow of SRAs

- SRA repair time = process time + AWP time. Installation of piece parts on SRAs is included in the process time.
- Process time is known and constant; AWP time is determined by the authorized stock of the piece parts used to fix the SRA, using a formula given below.

Inflow of piece parts

- The OST for each piece part is known and constant and may be different for different parts.

The assumptions are captured in Table 2.[1]

Section 4.2.1 gives formulas under the following additional assumptions:

- Piece parts are ordered under an (S – 1, S) policy;
- Repairs of NRFI SRAs begin immediately upon their removal from the WRA.

Section 4.2.2 gives formulas under an (S – s, S) policy and batching of SRA repairs.

Table 2

NU Assumptions

	NU	SU	U
<u>Outflow</u>			
WRA job arrival	D_k per day		
demands for SRA i	one every $1/D_k p_i$ days		
demands for piece part j going on WRA	one every $1/D_k p_j$ days		
demands for piece part j going on SRA i	one every $1/D_k q_j p_i$ days		
<u>Inflow of SRAs</u>			
process time	known and constant		
AWP time	deterministic		
<u>Inflow of piece parts</u>			
OST for piece part j	known and constant		

[1]They yield a D/D/n/∞/FIFO queue for each individual piece part or SRA not repaired locally.

4.2.1. Expected Waiting Times Under (S – 1, S) Ordering and No Batching of SRA Repairs

We first show how to compute average AWP time for SRAs; then, we show how to compute average AWP time for each WRA.

Average AWP time for SRAs. We begin by defining some notation and then we give the formulas. Let the SRAs that go on the WRA be indexed by i; we will consider a generic SRA labelled i. Let the piece parts that go on SRA i be indexed by j; at first, we will consider a generic piece part labeled j, and then we will consider all of the piece parts together. Let the RF of SRA i be p_i, and let the RF of piece part j—the probability that it is demanded by a job repairing SRA i—be q_j. Let n_j be the authorized stock level of piece part j, and let OST_j be the OST for piece part j. Define the *zero-wait quantity*

$$
\begin{aligned}
z_j &= [p_i \times q_j \times D_k \times OST_j] + 1 && \text{if } p_i \times q_j \times D_k \times OST_j \text{ is not an integer, and} \\
&= p_i \times q_j \times D_k \times OST_j && \text{if } p_i \times q_j \times D_k \times OST_j \text{ is an integer,}
\end{aligned}
$$

where [x] means "the integer part of x." Then the average time that SRA i waits for piece part j, if it needs it, is

$$W_j = \max\{0, OST_j - n_j/(p_i q_j D_k)\}, \tag{4.2.1}$$

which is positive when $n_j < z_j$ and 0 when $n_j \geq z_j$. To compute the average AWP time for SRA i, apply the tall-pole formula using W_j and q_j:

$$\text{average AWP for SRA i} = \sum_{j=1}^{m} W_j q_j \prod_{h<j} (1 - q_h), \tag{4.2.2}$$

where m is the number of distinct items on SRA i.

Note that the tall-pole formula has a different interpretation here than in Section 2, because under the NU assumptions there is no uncertainty and hence no probability. The meaning of the tall-pole formula here is as follows. Imagine a long sequence of SRA jobs. Piece part demands occur deterministically and evenly spaced in this sequence of SRA jobs, but at different rates, so each SRA job in the sequence has a known list of needed parts, but different SRA jobs have different lists. The tall-pole formula averages over the sequence of SRA jobs. This is why we said "average AWP time" above instead of "expected AWP time."

Actually, this not exact either, because of the discreteness of the NU assumptions. If an item has an RF of 0.3, it is needed in 3 of every 10 jobs, but 3 demands cannot be spaced evenly in 10 jobs. However, if the demands are placed in the sequence of SRA jobs in a regular pattern that has the right average frequency, the tall-pole approximation still holds.

Average AWP time for WRAs. We first give the formula for the average time the WRA waits for each of its parts if they are needed; then, we show how to compute the average AWP time for the WRA.

Two kinds of parts go on WRAs: piece parts and SRAs. The average amount of time the WRA waits for each of its piece parts is computed by a formula like the one given above for piece parts on SRAs. The average time the WRA waits for piece part j, if it needs it, is

$$W_j = \max\{0, OST_j - n_j/(p_j D_k)\}, \tag{4.2.3}$$

which is positive when $n_j < z_j$ and zero when $n_j \geq z_j$, where

$$
\begin{aligned}
z_j &= [p_j \times D_k \times OST_j] + 1 && \text{if } p_j \times D_k \times OST_j \text{ is not an integer, and} \\
&= p_j \times D_k \times OST_j && \text{if } p_j \times D_k \times OST_j \text{ is an integer.}
\end{aligned}
$$

To get W_i, the expected amount of time the WRA waits for SRA i, replace OST by {SRA i's process time} + {average AWP for SRA i} in the formula just given. This formula for the expected time the WRA waits for SRA i introduces a further approximation, which is discussed in Appendix A.

We have expected waiting times W_j for the parts that go on the WRA, and plugging them and the corresponding RFs into the tall-pole formula yields the average amount of time the WRA waits for parts.

4.2.2. Expected Waiting Times Under (S – s, S) Ordering and Batching of SRA Repairs

With given stock levels, a job waits longer for parts under (S – s, S) ordering than under (S – 1, S) ordering, because s demands instead of 1 must occur between orders. The average time the WRA waits for piece part j going directly on it is

$$
\begin{aligned}
W_j &= \frac{r - n + s}{s}\left\{OST - \frac{n - s + r + 1}{2pD}\right\} && \text{for } r - n + s > 0, \\
&= 0 && \text{for } r - n + s \leq 0,
\end{aligned}
\qquad (4.2.4)
$$

where the subscripts on n, D, and p have been suppressed and $r = \min(n_j, [p_j \times OST_j \times D_k])$. The zero-wait quantity under (S – s, S) ordering is r + s. For piece part j going on SRA i, replace pD in the above formula with $p_i q_j D_k$. If SRA i is repaired in batches of s, replace OST in the above formula with {SRA i's process time} + {average AWP for SRA i}. This representation of batching assumes that s batched SRAs can be repaired in the same time as a single SRA. If this is not true, an appropriate increment must be added to {SRA i's process time} + {average AWP for SRA i}.

The other stages of the calculation described in Section 4.2.1 are unchanged.

4.3. FORMULAS FOR EXPECTED WAITING TIMES UNDER SU ASSUMPTIONS

The SU assumptions differ from the NU assumptions in only one respect: in SU, the parts needed to fix an NRFI WRA or SRA are not known until the repair job begins, while under NU, timing of all part demands is known in advance. The SU assumptions are presented below and compared with NU in Table 3.

Outflow of parts:

- Inductions of WRA k occur equally spaced in time at a known rate of D_k per day, so the time between inductions is $1/D_k$ days.
- For each WRA job, a demand for SRA i occurs with probability p_i, the RF, independent of other parts and jobs.
- For each WRA job, a demand for piece part j (which goes directly on the WRA) occurs with probability p_j, the RF, independent of other parts and jobs.

- For each SRA job, a demand for piece part j used to fix the SRA, occurs with probability q_j, the RF, independent of other parts and jobs.

Inflow of SRAs:

- SRA repair time = process time + AWP time. Installation of piece parts on SRAs is included in the process time.
- Process time is known and constant; AWP time is determined by the authorized stock of the piece parts used to fix the SRA, using a formula given below.

Inflow of piece parts:

- The OST for each piece part is known and constant and may be different for different parts.

SU allows no uncertainty in the WRA induction stream, in OST, or in the process time of SRAs, but it allows uncertainty about which WRA and SRA jobs will require which parts. The NU and SU assumptions are summarized in Table 3.[2]

Section 4.3.1 gives formulas under the following additional assumptions:

- Piece parts are ordered under an (S – 1, S) policy;
- Repairs of NRFI SRAs begin immediately upon their removal from the WRA.

Section 4.3.2 gives formulas under an (S – s, S) policy and batching of SRA repairs.

Table 3

NU and SU Assumptions

	NU	SU	U
Outflow			
WRA job arrival	D_k per day	D_k per day	
demands for SRA i	one every $1/D_k p_i$ days	for each WRA job, Pr(demand) = p_i	
demands for piece part j going on WRA	one every $1/D_k p_j$days	for each WRA job, Pr(demand) = p_j	
demands for piece part j going on SRA i	one every $1/D_k q_j p_i$ days	for each SRA job, Pr(demand) = q_j	
Inflow of SRAs			
process time	known and constant	known and constant	
AWP time	deterministic	stochastic	
Inflow of piece parts			
OST for piece part j	known and constant	known and constant	

[2]The SU assumptions yield a F/D/n/∞/FIFO queue for each individual part, where the distribution F is one in which part demands can occur only on evenly spaced days, but the occurrence of demands is determined by a Bernoulli random variable.

4.3.1. Expected Waiting Times Under (S – 1, S) Ordering and No Batching of SRA Repairs

As in Section 4.2.1, we first show how to compute expected AWP time for SRAs; then, we show how to compute expected AWP time for WRAs.

Expected AWP time for SRAs. At first, we consider piece part j going on SRA i, and then we consider all the piece parts together. We use the notation defined in Section 4.2, with one change. For piece part j going on SRA i, define $z_j = [D_k \times OST_j]$. Then the expected amount of time that SRA i waits for piece part j, if it needs it, is

$$W_j = \frac{(p_i q_j)^n}{D} \sum_{r=n}^{z} (OST \times D - r) \binom{r-1}{n-1} (1 - p_i q_j)^{r-n}, \qquad (4.3.1)$$

where the subscripts on z, D, and n have been suppressed.

The expected AWP time for SRA i is computed using the tall-pole formula with W_j and q_j:

$$E\,(\text{AWP for SRA i}) = \sum_{j=1}^{m} W_j q_j \prod_{h<j} (1 - q_h), \qquad (4.3.2)$$

where m is the number of distinct items on SRA i.

The zero-wait quantity for SU is z_j if OST × D_k is an integer, and z_j + 1 otherwise. This is larger than the zero-wait quantity under NU by a factor of approximately $1/(p_i q_j)$. If either p_i or q_j is small, this factor can be substantial.

Expected AWP time for WRAs. As before, we first give the formula for the expected time that the WRA waits for each of its parts if they are needed; then, we say how to compute the expected AWP time for the WRA.

The expected time that the WRA waits for piece part j going directly on it is computed by a formula like the one for piece parts going on SRAs. The expected time the WRA waits for piece part j, if it needs it, is

$$W_j = \frac{p_j^n}{D} \sum_{r=n}^{z} (OST \times D - r) \binom{r-1}{n-1} (1 - p_j)^{r-n}, \qquad (4.3.3)$$

where the subscripts on z, D, and n have been suppressed. To get W_i, the expected amount of time the WRA waits for SRA i, replace OST by

{SRA i's process time} + E(AWP time for SRA i)

in the formula just given. This uses the approximation that was introduced in Section 4.2.

We now have expected waiting times W_j for all the parts that go on the WRA; plugging them and the RFs into the tall-pole formula yields the expected amount of time the WRA waits for parts.

4.3.2. Expected Waiting Times Under (S – s, S) Ordering and Batching of SRA Repairs

The expected time the WRA waits for piece part j depends on whether $s \leq n_j$ or $s > n_j$. If $s \leq n_j$, the expected time the WRA waits, if it needs piece part j, is

$$W_j = \frac{1}{s}\sum_{h=n-s+1}^{n}\left\{\sum_{r=h}^{z}\left(OST - \frac{r}{D}\right)\binom{r-1}{h-1}p^h(1-p)^{r-h}\right\}, \tag{4.3.4}$$

where the subscripts have been suppressed on n, D, and p. If $s > n_j$, the expected time the WRA waits is

$$W_j = \frac{1}{s}\left\{\sum_{h=1}^{n}\left[\sum_{r=h}^{z}\binom{r-1}{h-1}p^h(1-p)^{r-h}\left(OST - \frac{r}{D}\right)\right] + (s-n)OST + \frac{(s-n)(s-n-1)}{2pD}\right\}, \tag{4.3.5}$$

where again subscripts on n, D, p, and OST have been suppressed. The wait is zero for $n \geq z_j + s$. For piece parts going on SRA i, replace p in the above formulas with $q_j p_i$. For SRA i, replace OST in the above formula with {SRA i's process time} + E(AWP for SRA i). This representation of batching assumes that s batched SRAs can be repaired in the same time as a single SRA. If this is not true, an appropriate increment must be added to {SRA i's process time} + E(AWP for SRA i).

The other stages of the calculation described in Section 4.3.1 are unchanged.

4.4. FORMULAS FOR EXPECTED WAITING TIMES UNDER U ASSUMPTIONS

The difference between U and the other sets of assumptions is that U allows everything to be uncertain, with a specific stochastic structure, as follows.

Outflow of parts

- Inductions of WRA k occur in a homogeneous Poisson process: inter-arrival times are independent and exponentially distributed with constant rate λ_k (mean inter-arrival time $1/\lambda_k$). (Appendix A indicates how to use other distributions for inter-arrival times.)
- For each WRA job, a demand for SRA i occurs with probability p_i, the RF, independent of other parts and jobs.
- For each WRA job, a demand for piece part j (which goes directly on the WRA) occurs with probability p_j, the RF, independent of other parts and jobs.
- For each SRA job, a demand for piece part j used to fix the SRA occurs with probability q_j, the RF, independent of other parts and jobs.

Inflow of SRAs:

- SRA repair time = process time + AWP time. Installation of piece parts on SRAs is included in the process time.
- Process time is stochastic with a known expected value.

- AWP time is determined by the authorized stock of the piece parts used to fix the SRA, using a formula given below.

Inflow of piece parts:

- OST for piece part j follows an exponential distribution with rate δ_j (mean OST is $1/\delta_j$), independent of everything else; these rates can vary from part to part.

The three sets of assumptions are summarized in Table 4.[3] More general specifications—e.g., dropping some of the distributional or independence assumptions—might be useful for simulations, but we do not discuss them here.

Section 4.4.1 gives formulas under the following additional assumptions:

- Piece parts are ordered under an (S – 1, S) policy;
- Repairs of NRFI SRAs begin immediately upon their removal from the WRA.

Section 4.4.2 gives formulas under an (S – s, S) policy and batching of SRA repairs.

4.4.1. Expected Waiting Times Under (S – 1, S) Ordering and No Batching of SRA Repairs

As before, we first show how to compute expected AWP time for SRAs; then, we show how to compute expected AWP time for WRAs.

Table 4

NU, SU, and U Assumptions

	NU	SU	U
<u>Outflow</u>			
WRA job arrival	D_k per day	D_k per day	Poisson process, rate of λ_k per day
demands for SRA i	one every $1/D_k p_i$ days	for each WRA job, Pr(demand) = p_i	for each WRA job, Pr(demand) = p_i
demands for piece part j going on WRA	one every $1/D_k p_j$days	for each WRA job, Pr(demand) = p_j	for each WRA job, Pr(demand) = p_j
demands for piece part j going on SRA i	one every $1/D_k q_j p_i$ days	for each SRA job, Pr(demand) = q_j	for each SRA job, Pr(demand) = q_j
<u>Inflow of SRAs</u>			
process time	known and constant	known and constant	stochastic with known expectation
AWP time	deterministic	stochastic	stochastic
<u>Inflow of piece parts</u>			
OST for piece part j	known and constant	known and constant	exponential distribution with mean $1/\delta_i$

[3]*The U assumptions yield a M/M/n/∞/FIFO queue for each individual part or SRA not repaired locally.*

Expected AWP time for SRAs. At first, we consider piece part j, and then we consider all the piece parts together. The expected time that SRA i waits for piece part j, if it needs it, is

$$W_j = \frac{1}{\delta_j}\left(\frac{p_i q_j \lambda_k}{p_i q_j \lambda_k + \delta_j}\right)^n, \tag{4.4.1}$$

where the subscript on n is suppressed. The expected wait is positive for all authorized stock levels, reflecting the possibility, however scant, that many jobs arrive, most of them need this piece part, and that OST is long. To compute the expected AWP time for SRA i, apply the tall-pole formula using W_j and q_j:

$$E(\text{AWP for SRA i}) = \sum_{j=1}^{m} W_j q_j \prod_{h<j} (1 - q_h), \tag{4.4.2}$$

where m is the number of distinct items on SRA i.

Expected AWP time for WRAs. As before, we first give the formula for the expected time the WRA waits for each of its parts if they are needed, then we give the formula for the expected AWP time for the WRA.

The expected time the WRA waits for piece part j, which goes directly on it, is

$$W_j = \frac{1}{\delta_j}\left(\frac{p_j \lambda_k}{p_j \lambda_k + \delta_j}\right)^n, \tag{4.4.3}$$

which is the formula just given, with $p_i q_j$ replaced by p_j. To get W_i, the expected amount of time the WRA waits for SRA i, replace δ_j by

1/({SRA i's process time} + E(AWP time for SRA i))

in the most recent formula. This uses the approximation that was introduced in Section 4.2.

We now have expected waiting times W_j for all the parts that go on the WRA; plugging them and the RFs into the tall-pole formula yields the expected amount of time the WRA waits for parts.

4.4.2. Expected Waiting Times Under (S – s, S) Ordering and Batching of SRA Repairs

The expected time the WRA waits for piece part j depends on whether $s \le n_j$ or $s > n_j$. If $s \le n_j$, the expected time the WRA waits is

$$W_j = \frac{1}{s\delta_j} Q^{n-s+1} \frac{1-Q^s}{1-Q} \quad \text{for} \quad Q = \frac{p_j \lambda_k}{p_j \lambda_k + \delta_j}, \tag{4.4.4}$$

and if $s > n_j$, the expected time the WRA waits is

$$W_j = \frac{1}{s\delta_j}\frac{1-Q^{n+1}}{1-Q} + \frac{1}{s}\left\{\frac{s-n-1}{\delta_j} + \frac{(s-n)(s-n-1)}{2p_j\lambda_k}\right\}, \tag{4.4.5}$$

where Q is as above, and the subscript on n has been suppressed. For piece part j going on SRA i, replace p_j in the above formula with p_iq_j. For SRA i, replace δ_j by

1/({SRA i's process time} + E(AWP time for SRA i)).

This representation of batching assumes that s batched SRAs can be repaired in the same time as a single SRA. If this is not true, an appropriate increment must be added to {SRA i's process time} + E(AWP for SRA i).

The other stages of the calculation described in Section 4.4.1 are unchanged.

5. FORMULAS FOR THE SHORT-RUN PROBLEM

Section 3 described our method for selecting short-run supply actions to maximize aircraft availability, omitting the formula for the expected number of units of a WRA in AWP status at time t. This section gives that formula for the assumptions discussed in Section 3. We assume that a depot uses (S – 1, S) ordering for piece parts and that repairs of NRFI SRAs begin as soon as they are removed from the WRA. The formula can be generalized to (S – s, S) ordering and batching of SRA repairs, but we do not do so here.

For concreteness, we use a particular WRA, the radar antenna. The first step is to connect E(number of antennas in AWP status at time t) to formulas for individual parts, which we do in Section 5.1. In Section 5.2, we give formulas for piece parts going directly on the WRA. In Section 5.3 we give formulas for SRAs and for piece parts going on SRAs. Appendix A derives the formulas, and Appendix B lists all of the assumptions underlying these formulas.

5.1. CONNECTING E (NUMBER OF ANTENNAS IN AWP AT TIME t) TO FORMULAS FOR INDIVIDUAL PARTS

Recall that R antennas are on the repair schedule. Define a random variable X_r taking the value 1 if the r^{th} antenna on the schedule is in AWP status at time t and 0 if it is not. Then

$$\text{E(number of antennas in AWP at time t)} = \sum_{r=1}^{R} E(X_r) = \sum_{r=1}^{R} \Pr(X_r = 1). \qquad (5.1.1)$$

The r^{th} antenna is in AWP at time t if one or more of its parts arrives after time t, so

$$\sum_{r=1}^{R} \Pr(X_r = 1) = \sum_{r=1}^{R} \Pr(\text{some part needed by the } r^{th} \text{ antenna arrives after time t})$$

$$= R - \sum_{r=1}^{R} \Pr(\text{all parts needed by the } r^{th} \text{ antenna arrive before time t})$$

$$= R - \sum_{r=1}^{R} \left\{ \prod_{j} \Pr(\text{part j doesn't keep the } r^{th} \text{ antenna waiting after time t}) \right\}, \qquad (5.1.2)$$

where the product is over parts going directly on the antenna. The last equality holds because WRA induction times are fixed and known and parts fail independently.

Part j does not keep the r^{th} antenna waiting after time t if either of two things happens: The r^{th} antenna does not need part j, or it does need part j and part j arrives in time. Thus, by the law of total probability,

$$\Pr(\text{part j doesn't keep the } r^{th} \text{ antenna waiting after time t})$$

$$= \Pr(\text{part j doesn't keep the } r^{th} \text{ antenna waiting after time t} \mid \text{doesn't need part j}) \times \Pr(\text{doesn't need part j})$$

+ Pr(part j doesn't keep the r^{th} antenna waiting after time t | does need part j)
 × Pr(does need part j)

= 1 × (1 − p_j)
+ Pr(part j doesn't keep the r^{th} antenna waiting after time t | does need part j) × p_j,
(5.1.3)

where, as in Section 4, p_j is part j's RF. Given the above, we need only formulas for Pr(part j doesn't keep the r^{th} antenna waiting after time t | does need part j). These are given in Sections 5.2 and 5.3.

5.2. FORMULAS FOR PIECE PARTS GOING DIRECTLY ON WRAs

First, we define some notation. Let n_j be the number of units of part j that are either on the shelf or due in. Let d_m, for m = 1, 2, . . . , n_j, be the due dates of the n_j units of part j; if a unit is on the shelf, it has $d_m = 0$. Assign labels to units on the shelf and due in so that $d_1 \le d_2 \le \ldots \le d_{n_j}$. We need to consider cases in which more than n_j units of part j are needed to fix the R antennas on the schedule, so we need the due dates of units that are not on the due-in schedule at time 0. Because we assume (S − 1, S) ordering, the due date of the $(n_j + 1)^{th}$ unit is OST_j days after the first demand for part j, the due date of the $(n_j + 2)^{th}$ unit is OST_j days after the second demand for part j, and so on. Demands are stochastic; thus, the due-dates for the $(n_j + 1)^{th}$ and higher units are stochastic. An example of due-dates for part j is in Table 5. In Table 5, the first two units are on the shelf and the next $n_j - 2$ are due in between days 3 and 10.

Finally, define the indicator function I{condition}, where I{condition} = 1 if "condition" is true and 0 if it is false. We use conditions of the form "$t > d_m$", so that $I\{t > d_m\} = 1$ if d_m is less than t and 0 otherwise.

We can now give the formula for Pr(part j doesn't keep the r^{th} antenna waiting after time t), where for convenience we have dropped the condition "the r^{th} antenna does need part j." If $r \le n_j$, Pr(part j doesn't keep the r^{th} antenna waiting after time t) =

$$\sum_{m=0}^{r-1} I\{t > d_{m+1}\}\binom{r-1}{m} p_i^m (1-p_i)^{r-1-m}. \qquad (5.2.1)$$

Table 5

Due Dates for Part j

m	d_m
1	0
2	0
3	3
. . .	. . .
n_j	10
$n_j + 1$	1st demand + OST_j
$n_j + 2$	2nd demand + OST_j

If $r > n_j$, Pr(part j doesn't keep the r^{th} antenna waiting after time t) =

$$\sum_{m=0}^{n-1} I\{t > d_{m+1}\}\binom{r-1}{m} p_i^m (1-p_i)^{r-1-m}$$

$$+ \sum_{m=n}^{r-1} p_i^m (1-p_i)^{r-1-m}\left\{\sum_{h=m-n+1}^{r-n} I\{t > t_h + \mathrm{OST}\}\binom{h-1}{m-n}\binom{r-h-1}{n-1}\right\}, \qquad (5.2.2)$$

where t_h is the induction date of the h^{th} antenna, the subscript on n is suppressed, and $\binom{x}{-1}$ is defined to be 1 if $x = -1$ and 0 if $x > -1$. Generalizing these formulas to the case of stochastic due dates and OST involves replacing $I\{\bullet\}$ with the analogous probability.

5.3. FORMULAS FOR SRAs

We index SRAs by j and use the same notation as in Section 5.2, with one exception. Consider the case in which more than n_j units of SRA j are needed to fix the R antennas on the schedule. Because we assume (S − 1, S) ordering—that is, as soon as an NRFI SRA is removed, it is inducted for repair—the due date of the $(n_j + 1)^{th}$ unit of SRA j is {process time of SRA j} + {AWP time of that unit} after the first demand for SRA j. The due date for the $(n_j + 2)^{th}$ and subsequent units is defined analogously. Let T_j denote the process time of SRA j.

For $r \le n_j$, the formula for Pr(SRA j doesn't keep the r^{th} antenna waiting after time t) is identical to expression (5.2.1). For $r > n_j$, the formula is more complicated: Pr(SRA j doesn't keep the r^{th} antenna waiting after time t) =

$$\sum_{m=0}^{n-1} I\{t > d_{m+1}\}\binom{r-1}{m} p_i^m (1-p_i)^{r-1-m}$$

$$+ \sum_{m=n}^{r-1} p_j^m (1-p_j)^{r-1-m}\left\{\sum_{h=m-n+1}^{r-n} \binom{h-1}{m-n}\binom{r-h-1}{n-1}\Omega(h, m)\right\}, \qquad (5.3.1)$$

where the subscript on n is suppressed,

$$\Omega(h, m) = \Pr(t > t_h + T_j + \mathrm{AWP}_h \mid (m - n_j + 1)^{th} \text{ demand for SRA j occurred on } h^{th} \text{ WRA job}), \qquad (5.3.2)$$

t_h is the induction time of the h^{th} antenna on the schedule, and AWP_h is the AWP time of the unit of SRA j removed from the h^{th} antenna.

The computational formula for $\Omega(h, m)$ is complicated, but the complication is more a matter of bookkeeping than computational difficulty, and we offer some suggestions for simplifying the bookkeeping. Now

$$\Pr(t > t_h + T_j + \mathrm{AWP}_h \mid (m - n_j + 1)^{th} \text{ demand for SRA j occurred on } h^{th} \text{ WRA job}) \qquad (5.3.3)$$

is the probability that all the parts needed to fix the broken unit of SRA j removed from the h^{th} antenna arrive before time $t - T_j$, given that $m - n_j$ demands for SRA j occurred in the first $h - 1$ antenna jobs. We approximate this by the following product:

$$\prod_i \left\{1 - q_i + q_i \Psi(h, m - n_j)\right\} \tag{5.3.4}$$

where the product is over all parts i going on the SRA, and $\Psi(h, m - n_j)$ = Pr(part i doesn't keep SRA j waiting | $(m - n_j + 1)^{th}$ demand for SRA j occurred on h^{th} WRA job, the current SRA needs part i). The approximation that yields this product is discussed in Appendix A. Let n_i = {number of units of part i on the shelf at time 0} + {number of units of part i due in at time 0}. The form of $\Psi(h, m - n_j)$ depends on whether $m - n_j < n_i$ or $m - n_j \geq n_i$. If $m - n_j < n_i$, then $\Psi(h, m - n_j)$ =

$$\sum_{a=0}^{m-n_j} I\{t - T_j > d_{i,a+1}\} \binom{m - n_j}{a} q_i^a (1 - q_i)^{m-n_j-a} \tag{5.3.5}$$

where $d_{i,a+1}$ is the due date of the $(a + 1)^{th}$ unit of part i. This does not depend on h. If $m - n_j \geq n_i$, then $\Psi(h, m - n_j)$ =

$$\sum_{a=0}^{n_i-1} I\{t - T_j > d_{i,a+1}\} \binom{m - n_j}{a} q_i^a (1 - q_i)^{m-n_j-a}$$
$$+ \sum_{a=n_i}^{m-n_j} q_i^a (1-q_i)^{m-n_j-a} \sum_{b=a-n_i+1}^{m-n_j-n_i+1} \frac{\binom{b-1}{a-n_i}\binom{m-n_j-b}{n_i-1} \Xi(\gamma_{ij}, h, m-n_j, b)}{\binom{h-1}{m-n_j}} \tag{5.3.6}$$

where $\Xi(\gamma_{ij}, h, m - n_j, b)$ =

$$\Xi(\gamma_{ij}, h, m-n_j, b) = \begin{cases} 0 & \text{if } b > \gamma_{ij}, \\ \binom{h-1}{m-n_j} & \text{if } \gamma_{ij} \geq h-1, \\ \sum_{z=\xi}^{\gamma_{ij}} \binom{\gamma_{ij}}{z}\binom{h-\gamma_{ij}-1}{m-n_j-z} & \text{if } b \leq \gamma_{ij} < h-1, \end{cases} \tag{5.3.7}$$

for $\xi = \max(b, \gamma_{ij} + m - n_j - h + 1)$ and γ_{ij} is the number of antenna jobs with induction times less than or equal to $t - T_j - OST_i$.

These expressions for $\Psi(h, m - n_j)$ are formidable, but each sum includes just a few terms. The following approach may help in coding these expressions. For each piece part i going on SRA j, construct a three-way array with levels corresponding to the index a and within each level, rows and columns corresponding to $m - n_j$ and h, respectively. For each value of a, then, the array gives a two-way table in $m - n_j$ and h, depicted in Figure 10.

RAND#544-10-0793

$m - n_j$	$h \rightarrow$	1	2	3	···	$R - n_j - 1$	$R - n_j$	Number of a entries
0		X	X	X	···	X	X	1
1			X	X	···	X	X	2
2				X		X	X	3
•						•	•	•
•						•	•	•
•						•	•	•
$R - n_j - 1$							X	$R - n_j$

Figure 10—Two-Way Array in $m - n_j$ and h

The two-way table in $m - n_j$ and h is the rectangle enclosed by double lines. The rows are labeled by values of $m - n_j$ and the columns by values of h. Cells containing an X contribute to $\Psi(h, m - n_j)$. Empty cells do not contribute to $\Psi(h, m - n_j)$. In the three-way array of which this is a level, the levels are labelled by $a = 0, 1, \ldots, R - n_j - 1$. The rightmost column, labeled "number of a entries," gives the number of values of a (i.e., levels of the three-way array in a, $m - n_j$ and h) that contribute to $\Psi(h, m - n_j)$, these levels being $a = 0, 1, \ldots, m - n_j$.

To fill the cells of the three-way array for part i, follow these rules. For levels $a < n_i$, the contents of the cell in the a^{th} level, $(m - n_j)^{th}$ row, h^{th} column is the a^{th} term of Expression (5.3.5). This contribution does not depend on h. For levels $a \geq n_i$, the contents of the cell in the a^{th} level, $(m - n_j)^{th}$ row, h^{th} column is the a^{th} term of Expression (5.3.6), which does depend on h.

Having filled the cells of the three-way array for part i, $\Psi(h, m - n_j)$ is computed as follows. For each part i, sum across the levels of a to yield a two-way table in $m - n_j$ and h. We now have two-way tables for each part i; stack them to form a three-way array with levels indexed by i. Multiply across the levels of i to yield a two-way table in $m - n_j$ and h. This table contains $\Psi(h, m - n_j)$.

The amount of computing required is not as extensive as it may seem. The table containing $\Psi(h, m - n_j)$ need not be recomputed after changes to the due-in schedule for SRA j. If dates of existing due-ins of SRA j are changed, only terms $m \leq n_j$ in Expressions (5.2.1) and (5.2.2) are affected. If new units of SRA j are added to the due-in schedule, increasing n_j, Ψ need not be recomputed because it is a function of $m - n_j$, not n_j. If a single part i on SRA j is affected by a supply action, the three-way array for that part must be recomputed but the three-way array for other parts on the SRA need not. When a part i is affected by a supply action, only some cells in its three-way array are affected, but it is probably simpler to recompute its entire three-way array.

6. CONCLUSION

This report describes methods for solving problems of investment and operations related to parts supply for Naval aviation depots. The presentation of the long-run method focuses on setting stock levels to minimize AWP times and the value of the repair pipeline, but the method can be used for many other purposes. For example, it can be used to examine the effects of reducing OSTs, as compared with the payoff available from other uses of the same amount of investment. Given the current size of OSTs and procurement lead times in the Navy, reducing OSTs probably has a larger payoff than more efficient stock levels.

In developing the methods, we made some assumptions that will restrict their usefulness. The most important of these are the assumptions that all part needs are known upon induction of a WRA and that all needed parts are needed immediately. While these assumptions will be innocuous for some classes of components, they will be a much greater concern for items involving substantial processing, such as engines or hydraulic equipment. Some have argued that our definition of AWP time is appropriate because it measures the extent to which parts are a constraint on repair and because improving on these assumptions would require enshrining current repair processes in our algorithm.

Elaborations of the methods presented here are possible at the cost of greater complication and computing and data requirements.

Appendix A

DERIVATIONS OF THE FORMULAS IN SECTIONS 4 AND 5

Sections A.1 through A.3 derive the formulas given in Section 4 for the long-run method. Section A.4 derives the formulas given in Section 5 for the short-run method.

A.1. FORMULAS FOR THE LONG-RUN PROBLEM, UNDER NU ASSUMPTIONS

A.1.1. Under (S – 1, S) Ordering and No Batching of SRA Repairs

We derive the time a WRA waits if it needs piece part j going directly on the WRA, Expression (4.2.3). We assume (S – 1, S) ordering, so each demand for a unit of part j triggers an order for a unit of part j. Thus, as suggested by Figures 4–6 in Section 2, the current demand for part j is filled by the unit ordered after the n^{th} previous demand (suppressing the subscript on n) and the wait for part j is

$$\max(0,\text{OST} - \{\text{time since the } n^{th} \text{ previous demand}\}), \tag{A.1.1}$$

where the subscript on OST has been suppressed. But a demand occurs every $1/p_jD_k$ days, so the time since the n^{th} previous demand is n/p_jD_k days. This yields expression (4.2.3). An analogous argument yields expression (4.2.1) for piece parts used to fix SRAs.

The formula for SRA i is derived similarly but uses an additional approximation. The time the WRA waits for SRA i, if it needs it, is

$$\max(0,\{\text{repair time of SRA i}\} - \{\text{time since the } n^{th} \text{ previous demand for SRA i}\}). \tag{A.1.2}$$

But {repair time of SRA i} = {process time of SRA i} + {AWP time of SRA i}, and because different units of SRA i need different repair parts, AWP time will vary across units of SRA i. Thus, the average time the WRA waits for SRA i is the average of expression (A.1.2) across jobs repairing SRA i. This latter can be approximated by moving the average operator inside the maximization operator, yielding the formula in Section 4.2.

A.1.2. Under (S – s, S) Ordering and Batching of SRA Repairs

We derive the formula for piece part j going directly on the WRA, Expression (4.2.4). Recall that if the authorized stock of part j is n units (suppressing the subscript), the current demand for part j is filled by the unit ordered after the n^{th} previous demand. Under (S – 1, S) ordering, that requisition was placed after the n^{th} previous demand; under (S – s, S) ordering, that requisition may not be placed until several more demands have occurred. If the $(n + 1)^{th}$ previous demand triggered an order, then the n^{th} previous demand was the first after the order, so the order replacing the n^{th} previous demand is placed after the $(n - s + 1)^{th}$ previous demand. If the $(n + 2)^{th}$ previous demand triggered an order, then the n^{th} previous demand was the second after the order, so the order replacing the n^{th} previous demand is placed after the $(n - s + 2)^{th}$ previous demand, etc. If $s > n$, then $n - s + m$ (for $0 < m \le s$) could be negative, in which case the order replacing the n^{th} previous demand is placed after the $(s - n - m)^{th}$ *future* demand, where the "0^{th}" future demand is the current demand. If we pick a WRA job at random, as these long-run calculations do, and assume that the job demands

part j, we do not know whether the n^{th} previous demand was the first, second, . . . , or s^{th} demand after the order that preceded it; all s possibilities are equally likely. Thus, the expected wait is

$$\sum_{m=1}^{s} (\text{wait} \mid n^{th} \text{ previous demand was the } m^{th} \text{ after the preceding order})$$

$$\times \Pr(n^{th} \text{ previous demand was the } m^{th} \text{ after the preceding order})$$

$$= \frac{1}{s} \sum_{m=1}^{s} (\text{wait} \mid n^{th} \text{ previous demand was the } m^{th} \text{ after the preceding order}).$$

The summand is OST – n/pD—the waiting time if s = 1—plus the amount of time needed for s – m further demands to occur, which is (s – m)/pD. Thus, the expected wait is

$$\frac{1}{s} \sum_{m=1}^{s} \max(0, \text{OST} - \frac{n}{pD} + \frac{s-m}{pD}). \qquad \text{(A.1.3)}$$

Define $r = \min(n, [p \times \text{OST} \times D])$, where [x] is "the integer part of x." The summand is positive for $m \le r - n + s$, and if this is used to get rid of the maximum operator in (A.1.3), some algebra shows (A.1.3) equal to

$$\frac{r-n+s}{s} \left\{ \text{OST} - \frac{n-s+r+1}{2pD} \right\}, \qquad \text{(A.1.4)}$$

when $r - n + s > 0$. When $r - n + s \le 0$, the sum in (A.1.3) is null and the expected wait is 0. This gives the zero-wait quantity r + s.

Analogous arguments yield expected waits for SRA i and for piece part j on SRA i.

A.2. FORMULAS FOR THE LONG-RUN PROBLEM, UNDER SU ASSUMPTIONS

A.2.1. Under (S – 1, S) Ordering and No Batching of SRA Repairs

We derive the formula for piece part j going directly on the WRA, Expression (4.3.3). Suppose the current WRA job needs a unit of part j. Because we assume (S – 1, S) ordering, the unit satisfying this demand was ordered after the n^{th} previous demand (again, suppressing subscripts). Under the SU assumptions, the timing of previous demands for part j is stochastic, but previous demands can occur only on dates on which WRA jobs are inducted. If the n^{th} previous demand for part j occurred on the r^{th} previous WRA job—where $r \ge n$ because it takes at least n previous jobs to accumulate n demands—then the current job waits

$$\max(0, \text{OST} - r/D) \qquad \text{(A.2.1)}$$

for its unit of part j, suppressing all subscripts. The n^{th} previous demand for part j occurs on the r^{th} previous WRA job if and only if

- A demand for part j occurs on the r^{th} previous job, and
- Exactly $n - 1$ demands occur in the next $r - 1$ jobs.

Thus, Pr(n^{th} previous demand for part j occurred on the r^{th} previous WRA job) =

$$\binom{r-1}{n-1} p^n (1-p)^{r-n}. \tag{A.2.2}$$

Thus, the WRA's expected wait for part j is

$$p^n \sum_{r=n}^{z} \left(OST - \frac{r}{D}\right) \binom{r-1}{n-1} (1-p)^{r-n}, \tag{A.2.3}$$

where the summation includes values of r for which $OST \geq r/D$. A simple manipulation yields Expression (4.3.3).

The formula for piece part j going on SRA i, Expression (4.3.1), follows from an analogous argument. The formula for SRA i follows by applying the approximation described in Section A.1.

A.2.2. Under (S – s, S) Ordering and Batching of SRA Repairs

We derive the formula for piece part j going directly on the WRA, Expressions (4.3.4) and (4.3.5). By the argument used in Section A.1.2, the expected wait is

$$\frac{1}{s} \sum_{m=1}^{s} E(\text{wait} \mid n^{th} \text{ previous demand was the } m^{th} \text{ after the preceding order}). \tag{A.2.4}$$

The sum in expression (A.2.4) is equal to

$$\sum_{h=n-s+1}^{n} E(\text{wait} \mid \text{the order replacing the } n^{th} \text{ previous demand is placed after the } h^{th} \text{ previous demand}). \tag{A.2.5}$$

If $n \geq s$, the summand in (A.2.5) is equal to

$$\sum_{r=h}^{z} E(\text{wait} \mid \text{the } h^{th} \text{ previous demand was on the } r^{th} \text{ previous job}) \times \Pr(\text{the } h^{th} \text{ previous demand was on the } r^{th} \text{ previous job}),$$

which, by the argument used in Section A.2.1, is equal to

$$\sum_{r=h}^{z} \left(OST - \frac{r}{D}\right)\binom{r-1}{h-1} p^h (1-p)^{r-h}. \tag{A.2.6}$$

Expression (4.3.4) follows immediately. If $n < s$, Expression (A.2.5) is still correct, but its interpretation changes. For $h \le 0$, the wait is for the order resulting from the $(-h)^{th}$ future demand, so the expected wait is

$$OST + E(\text{wait until } (-h)^{th} \text{ future demand}), \tag{A.2.7}$$

and $h = 0$ corresponds to the current demand. The time between WRA jobs is 1/D, and a simple probabilistic argument shows that the expected number of WRA jobs between demands for part j is 1/p, so Expression (A.2.7) equals OST + h/pD. For $h > 0$, Expression (A.2.6) still holds, and Expression (4.3.5) follows immediately.

The derivations for piece part j going on SRA i and for SRA i are analogous.

A.3. FORMULAS FOR THE LONG-RUN PROBLEM, UNDER U ASSUMPTIONS

A.3.1. Under (S – 1, S) Ordering and No Batching of SRA Repairs

We give the derivation for piece part j going directly on the WRA, Expression (4.4.3). Suppose the current WRA job needs part j. Now

$$E(\text{wait for part } j) = 0 \times Pr(\text{wait} = 0) + E(\text{wait} \mid \text{wait} > 0)\, Pr(\text{wait} > 0),$$

so we need to derive $E(\text{wait} \mid \text{wait} > 0)$ and $Pr(\text{wait} > 0)$.

First, to derive $E(\text{wait} \mid \text{wait} > 0)$. If part j has authorized stock of n units, the unit that fills the current demand was ordered after the n^{th} previous demand for part j. Let T be the time back to the n^{th} previous demand. The waiting time is the excess of OST—which has an exponential distribution with rate δ_j—over T, which has probability density f(T), which we will specify later. Therefore

$$\begin{aligned}
&E(\text{wait} \mid \text{wait} > 0) \\
&= \int_0^\infty E(\text{wait} \mid \text{wait} > 0, T)\, f(T)\, dT \\
&= \int_0^\infty E(OST - T \mid OST > T, T)\, f(T)\, dT \qquad \text{(A.3.1)} \\
&= \int_0^\infty \frac{1}{\delta_j} f(T)\, dT, \text{ by the memoryless property of the exponential distribution} \\
&= 1/\delta_j, \text{ for all densities } f(T).
\end{aligned}$$

Now we derive Pr(wait > 0). By a similar argument,

$$
\begin{aligned}
\Pr(\text{wait} > 0) &= \Pr(\text{OST} > T) \\
&= \int_0^\infty \Pr(\text{OST} > T \mid T)\, f(T)\, dT \\
&= \int_0^\infty \exp(-\delta_j T)\, f(T)\, dT
\end{aligned}
\qquad \text{(A.3.2)}
$$

because OST follows an exponential distribution. Expression (A.3.2) is the characteristic function of T evaluated at $i\delta_j$, where i is the solution to $i^2 = -1$. T is the sum of the n times between the n last demands, and by a standard result in probability theory, the characteristic function of the sum of independent random variables is the product of the characteristic functions of the summands. Thus, Pr(wait > 0) can be obtained for arbitrary distributions for the times between inductions of WRAs. (Presumably one would want distributions with closed-form characteristic functions.)

In Section 4.4, we assumed that times between WRA inductions had an exponential distribution with rate λ (suppressing the subscript), so T has density function

$$f(T) = \frac{(p_j\lambda)^n T^{n-1}}{\Gamma(n)} \exp(-\, p_j \lambda T)$$

for $T > 0$, where $\Gamma(n)$ is the gamma function evaluated at n. This standard result can be obtained by summing n exponential random variables with rates $p_j\lambda$. Some calculus yields expression (4.4.3).

It is easy to change the assumption that inter-arrival times for WRA jobs are exponential but much more difficult to change the assumption that OST has an exponential distribution. Doing so changes Equations (A.3.1) and (A.3.2), generally yielding more complicated integrals. Also, if OST does not have an exponential distribution, a substantial restriction is placed on the class of distributions of T for which Pr(wait > 0) has a closed form.

The argument for the formula for piece part j going on SRA i is almost identical. For SRA i, the expression given in section 4.4 is obtained using the approximation introduced in Section A.1.

A.3.2. Under (S – s, S) Ordering and Batching of SRA Repairs

Expressions (4.4.4) and (4.4.5) follow by an argument analogous to the one given in Section A.2.2.

A.4. FORMULAS FOR THE SHORT-RUN PROBLEM

This section derives formulas for Pr(part j doesn't keep the r^{th} antenna waiting after time t | the r^{th} antenna needs part j). Section A.4.1 derives Expressions (5.2.1) and (5.2.2) for piece parts going directly on the antenna. Section A.4.2 derives Expression (5.3.1) for SRAs going on the antenna and derives subsidiary formulas for piece parts going on SRAs.

A.4.1. Piece Parts Going Directly on WRAs

Consider a piece part j, let Z be the event "part j doesn't keep the r^{th} antenna waiting after time t," and suppress the conditioning event "the r^{th} antenna needs part j" for convenience. We need a formula for Pr(Z). By the law of total probability, Pr(Z) =

$$\sum_{m=0}^{r-1} \Pr(Z \mid \text{m demands for part j in first r} - 1 \text{ jobs})$$
$$\times \Pr(\text{m demands for part j in first r} - 1 \text{ jobs}). \qquad \text{(A.4.1)}$$

Pr(m demands for part j in first r – 1 jobs) is $\binom{r-1}{m} p_i^m (1-p_i)^{r-1-m}$, the binomial distribution. Deriving Pr(Z | m demands for part j in first r – 1 jobs) is more complicated. If r $\le n_j$, then the current demand for part j will necessarily be filled by a unit that was either on the shelf or due in at time 0. If there were m demands for part j in the first r – 1 antenna jobs, the current demand will be filled by the $(m + 1)^{th}$ unit, on day d_{m+1}. In that case, the event Z occurs if $d_{m+1} < t$ and does not occur if $d_{m+1} > t$, so Pr(Z | m demands for part j in first r – 1 jobs) = $I\{t > d_{m+1}\}$, and expression (5.2.1) follows immediately.

If $r > n_j$ but m, the number of previous demands for part j, is less than n_j, the current demand will be filled by a unit that was either on the shelf or due in at time 0, so the argument in the last paragraph yields Pr(Z | m demands for part j in first r – 1 jobs) = $I\{t > d_{m+1}\}$ for $m < n_j$. If $r > n_j$ and $m \ge n_j$, the current demand for part j is filled by the unit ordered after the $(m - n_j + 1)^{th}$ demand for part j. Because the parts needed to fix each antenna are stochastic, the $(m - n_j + 1)^{th}$ demand for part j occurs stochastically and the part that will fill the current demand arrives OST days later. By another application of the law of total probability (and using abbreviated notation), Pr(Z | m previous demands) =

$$\sum_{h=1}^{r-1} \Pr(Z \mid \text{m previous demands, } (m-n+1)^{th} \text{ demand was on } h^{th} \text{ antenna job})$$
$$\times \Pr(\text{m previous demands, } (m-n+1)^{th} \text{ demand was on } h^{th} \text{antenna job}), \qquad \text{(A.4.2)}$$

where the subscript on n has been suppressed. As in the earlier arguments,

$$\Pr(Z \mid \text{m previous demands, } (m-n+1)^{th} \text{ demand was on } h^{th} \text{ antenna job})$$
$$= I\{t > OST + t_h\},$$

where t_h is the induction time of the h^{th} antenna job. The event "m previous demands, $(m - n + 1)^{th}$ demand was on h^{th} antenna job" occurs if and only if

- $m - n_j$ demands for part j occur in the first h – 1 antenna jobs,
- a demand for part j occurs on the h^{th} antenna job, and
- $n_j - 1$ demands occur in the r – h – 1 jobs between the h^{th} job and the $(r - 1)^{th}$ job.

This yields Expression (5.2.2). The limits on h in the summation in that expression follow from the need to have $h - 1 \ge m - n_j$ and $r - h - 1 \ge n_j - 1$. If $n_j = 0$, this derivation works if

$\binom{x}{-1}$ is defined to be 0 if $x > -1$ and 1 if $x = -1$. (If the binomial coefficient is expressed as a ratio of gamma functions, this convention follows by a limiting argument.)

A.4.2. SRAs and Piece Parts Going on SRAs

The derivation of Pr(Z)—that is, Pr(SRA j doesn't keep the r^{th} antenna waiting after time t | the r^{th} antenna needs SRA j)—is identical to the derivation for piece parts, up to the derivation of Expression (5.2.1) covering the case $r \le n_j$. For $r > n_j$, the argument for piece parts holds for SRAs except that Pr(Z | m previous demands for SRA j, $(m - n + 1)^{th}$ demand was on h^{th} antenna job) is now determined by the completion of the repair of the NRFI SRA that was removed from the h^{th} antenna. But the event

"SRA j doesn't keep the r^{th} antenna waiting after time t | m previous demands for SRA j, $(m - n_j + 1)^{th}$ demand was on h^{th} antenna job"

is the same as the event

"$t > t_h + AWP_h + T_j$ | m previous demands for SRA j, $(m - n_j + 1)^{th}$ demand was on h^{th} antenna job,"

where t_h is as above, T_j is the process time of SRA j, and AWP_h is the AWP time of the unit of SRA j removed from the h^{th} antenna. These two events are the same because the unit of SRA j removed from the h^{th} antenna becomes available to the r^{th} antenna at $t_h + AWP_h + T_j$: t_h is the time it is removed from the h^{th} antenna and enters repair, AWP_h is the time it waits for parts, and T_j is its processing time. The equivalence of these two events yields Expression (5.3.1).

Now to derive $\Omega(h, m)$. Abbreviating the notation slightly, $\Omega(h, m)$ =

$$\Pr(t > t_h + AWP_h + T_j \mid \text{m previous demands, } (m - n_j + 1)^{th} \text{ demand was on } h^{th} \text{ antenna job}) \quad (A.4.3)$$

This is equal to

$$\Pr(\text{all parts arrive before } t - T_j \mid \text{m previous demands, } (m - n_j + 1)^{th} \text{ demand was on } h^{th} \text{antenna job})$$

$$= \Pr(\text{last part arrives before } t - T_j \mid \text{m previous demands, } (m - n_j + 1)^{th} \text{ demand was on } h^{th} \text{ antenna job}).$$

$$= \sum_C \Pr(\text{last part arrives before } t - T_j \mid Q, C) \Pr(C \mid Q) \quad (A.4.4)$$

where Q is the conditioning event "m previous demands, $(m - n_j + 1)^{th}$ demand was on h^{th} antenna job," C is a configuration of $m - n_j$ demands for SRA j in the first $h - 1$ antenna jobs, and the sum is over all such configurations C. We were able to avoid this last sum for piece parts going directly on WRAs (in Section A.4.1) because we knew the antenna induction times, but here we do not know when the SRA inductions occur; SRAs are inducted

immediately after being removed from antennas, and we do not know which antennas will have NRFI SRAs. However, given a configuration C of $m - n_j$ demands for SRA j in the first $h - 1$ antenna jobs, it follows that

$$\sum_C \Pr(\text{last part arrives before } t - T_j \mid Q, C)\Pr(C \mid Q)$$

$$= \frac{1}{\binom{h-1}{m-n_j}} \sum_C \left\{ \prod_i \Pr(\text{part i doesn't keep the SRA waiting past } t - T_j \mid Q, C) \right\} \quad \text{(A.4.5)}$$

where the product is over all parts on the SRA. This result follows because given Q and C, demands for parts to repair SRA j are independent and all $\binom{h-1}{m-n_j}$ configurations C of $m - n_j$ demands for SRA j in the first $h - 1$ antenna jobs are equally likely. This expression for Ω (h, m) is exact. Our approximation follows if the summation over C and the product across i are exchanged, and $\Psi(h, m - n_j)$ is defined to be

$$\frac{1}{\binom{h-1}{m-n_j}} \sum_C \Pr(\text{part i doesn't keep the SRA waiting past } t - T_j \mid Q, C, \text{SRA needs part i}).$$

(A.4.6)

This approximation is exact if Pr(part i doesn't keep the SRA waiting past $t - T_j \mid Q,C$) does not depend on C. As will be seen below, this occurs if $m - n_j < n_i$, i.e., all parts used to fix the SRA are drawn from stocks on the shelf or in the due-in pipeline at time 0.

We now need to derive Expressions (5.3.5) and (5.3.6), the computing formulas for $\Psi(h, m - n_j)$. Recall that we are considering whether the unit of SRA j removed from the h^{th} antenna waits for piece parts past $t - T_j$. This unit will satisfy the $(m + 1)^{th}$ demand for SRA j; n_j demands for SRA j were satisfied off the shelf or from units due in at time 0, so $m - n_j$ units were repaired previously. Thus, derivation of $\Psi(h, m - n_j)$ is similar to the derivation in Section A.4.1. Define the event Z_i to be "the SRA removed from the h^{th} antenna is not delayed past $t - T_j$ by part i | C, Q, and that SRA needs part i"; we want $\Pr(Z_i)$. By the argument in Section A.4.1, if $m - n_j < n_i$—that is, the number of units of part i on the shelf or due in at time 0 exceeds the number of previous jobs repairing SRA j—then the demand for part i for the h^{th} job will necessarily be satisfied by a unit of part i that was either on the shelf or due in at time 0. Therefore, the availability date of that unit depends only on the number of previous demands for part i, not on their timing; so it does not depend on the configuration C of previous demands for SRA j, only on their number. Expression (5.3.5) follows immediately.

Now suppose $m - n_j \geq n_i$. If a, the number of previous demands for part i, is less than n_i, the current demand for part i will be filled by a unit that was either on the shelf or due in at time 0, so the argument used in the last paragraph yields the first summation in Expression (5.3.6). That leaves $m - n_j \geq n_i$ and $a \geq n_i$; by the argument used before, the current demand for part i is filled by the unit ordered after the $(a - n_i + 1)^{th}$ demand for part

i. The $(a - n_i + 1)^{th}$ demand for part i occurred stochastically in one of the first $m - n_j$ jobs repairing SRA j, so its timing depends on when those $m - n_j$ jobs commenced—which, by assumption, are the times that those NRFI SRAs were removed from their antennas. Thus, the second term of Expression (5.3.6) depends on the configuration C of the first $m - n_j$ demands for SRA j in the first $h - 1$ antenna jobs. Fix the configuration C and apply the argument that was used to derive Expressions (5.2.2) and (5.3.1), using b as the index of summation, analogous to h in Expressions (5.2.2) and (5.3.1). Reversing the order of summation of a and b yields Expression (5.3.6), where $\Xi(\gamma_{ij}, h, m - n_j, b)$ is the number of configurations C for which the b^{th} SRA demand (generating the $(a - n_i + 1)^{th}$ demand for part i) occurred before $t - T_j - OST_i$.

Finally, we derive the expression for $\Xi(\gamma_{ij}, h, m - n_j, b)$, which is the number of ways of placing $m - n_j$ SRA demands in $h - 1$ antenna jobs, in such a way that b of the SRA demands occur before $t - T_j - OST_i$. Let γ_{ij} be the number of antenna jobs with induction times less than or equal to $t - T_j - OST_i$. We need the number of ways of placing $m - n_j$ SRA demands in $h - 1$ antenna jobs so that the b^{th} demand occurs no later than the ${\gamma_{ij}}^{th}$ antenna job. If $b > \gamma_{ij}$, clearly $\Xi = 0$. If $\gamma_{ij} > h - 1$, then all $\binom{h-1}{m-n_j}$ ways to place the SRA demands have the b^{th} demand occurring before the ${\gamma_{ij}}^{th}$ antenna job. So assume $b \le \gamma_{ij} \le h - 1$. Then Ξ is the number of ways to place at least b SRA demands in the first γ_{ij} antenna jobs, with the remaining demands being placed in the last $h - \gamma_{ij} - 1$ antenna jobs. This yields Expression (5.3.7).

Probabilistic OST changes the derivation of Ξ. Instead of being the number of ways of placing $m - n_j$ SRA demands in $h - 1$ antenna jobs, in such a way that b of the SRA demands occur before $t - T_j - OST_i$, Ξ is $\sum \Pr(t - T_j > OST_i + \tau_b)$, where the sum is over configurations C and τ_b is the time of the b^{th} demand for SRA j.

Appendix B

ASSUMPTIONS UNDERLYING THE LONG-RUN AND SHORT-RUN METHODS

Section B.1 presents the assumptions of the long-run method and discusses what we know of their effects. Section B.2 presents the assumptions of the short-run method. Because we have not tested the short-run method, our discussion of the effects of the assumptions is short.

B.1. ASSUMPTIONS FOR THE LONG-RUN PROBLEM

B.1.1. List of Assumptions

The following list categorizes the assumptions for the long-run method according to their function.

Definitions:

- The depot is at arm's length financially with all its suppliers.
- Time homogeneity (steady-state).
- To define AWP time:
  - — all part needs are known at induction
  - — all needed parts are needed at induction.
- Cost is the sum of the unit prices of authorized stock.
- WRA and SRA turnaround time = AWP + process time, with part installation time included in process time.
- SRA repairs require only piece parts.

For ease of computation:

- The tall-pole approximation.

Data availability:

- Accurate bills-of-material, including RFs for piece parts and SRAs.

To facilitate inventory theory:

- The depot has adequate repair capacity (infinite servers).
- The induction rate of a WRA is independent of the number of units in the repair pipeline.
- No condemnations of SRAs or WRAs.
- No common items.
- No cannibalization.
- FIFO (service in order of arrival).
- WRA inductions and OST are stochastically independent.
- UPA is 1.
- (S – s, S) ordering.

- Induction of WRA jobs:
  - NU and SU: evenly spaced and deterministic
  - U: homogeneous Poisson process.
- OST from the supply system:
  - NU and SU: constant and deterministic
  - U: exponential distribution.
- Demands for parts to fix WRAs and SRAs:
  - NU: a deterministic fraction of WRA or SRA jobs
  - SU and U: Bernoulli, where the probability a part will be demanded in a job is its RF and parts are independent of one another.

B.1.2. Discussion of the Effects of the Assumptions

As noted in the main body of the report, perhaps the most important assumptions are that all part needs are known upon induction and that all needed parts are needed immediately. The second assumption clearly increases AWP time: if a part is not needed until day 10 and arrives on day 9, it will contribute nothing to AWP time in reality but will contribute 9 days in our computation. The effect of the first assumption is to decrease AWP time: if a part takes 10 days to deliver and it is ordered on day 1, it will arrive sooner than if the need for it were discovered on day 5, it were ordered then, and it took 10 days to deliver. We have not been able to test the effect of these two assumptions, but we believe it is the most important potential defect of the long-run method.

The cost assumptions are discussed in Section 2. Except for systems that are just moving to organic depot repair, the cost assumptions will tend to understate the cost of authorized stock and overstate the benefits. Nonetheless, as an index for comparing parts, the rate of return should still be appropriate.

We do not know the effect of inaccuracy in the bills-of-material or RFs.

The "no cannibalization assumption" will tend to increase the value of cannibalizable parts. Although we have not done so, it should be straightforward to permit cannibalization, by this device: if n parts on a WRA cannot be cannibalized and m parts can, the m parts can be treated as a single super part, which is available when one unit of each of the m parts is available. Permitting condemnations should also be straightforward. (S − s, S) ordering seems innocuous. We have worked around the "no common items" assumption by creating dummy items for each WRA a common item is used to fix. The effect of this workaround is unclear. On one hand, it could increase the value of a common item (and thus increase its authorized stock) because its units cannot be used as flexibly as they would be in reality. On the other hand, it could decrease a common item's value because each WRA the common item is used to fix has a lower demand rate than do all of those WRAs combined.

The other miscellaneous assumptions will be hard to avoid, and their effects are unclear.

The tall-pole formula and assumptions about induction of WRA jobs and OST. Two things make the tall-pole formula difficult to avoid. First, for the U assumptions, the stochastic timing of WRA inductions induces a correlation in the timing of demands for parts. Thus, waiting times for individual parts are correlated, which complicates the derivation of the expected value of the maximum waiting time. For the SU assumption, this problem does not affect piece parts or SRAs going on WRAs, because the WRAs are inducted at fixed, known times, but it does affect piece parts used to fix SRAs. The second problem is that even if waiting times for parts were not correlated, the exact expected value of the maximum

waiting time is easy to write down but takes a long time to compute. Without the tall-pole formula, we saw no practicable way to proceed.

There are two distinct concerns:

- How much does the tall-pole formula understate expected AWP time?
- How much does it affect evaluation of investments?

We have paper-tested the effect of the tall-pole formula and the assumptions about WRA inductions and OST. This section describes those tests, the results of which are quite positive, as far as they go.

How much does the tall-pole formula understate expected AWP time? The tall-pole formula approximates $E(\max\{X_1, X_2, \ldots, X_m\})$ by $\max\{E(X_1), E(X_2), \ldots, E(X_m)\}$, where X_i is 0 if part i is not needed and is part i's stochastic waiting time if it is needed, and m is the number of items on an SRA or WRA. It is easy to show that the tall-pole formula is never greater than the exact formula and that it understates E(AWP) more for larger m or as the variances of the Xs increase. Apart from this, we have little specific information about how the tall-pole formula understates expected AWP time.

We conjecture that this understatement has an important consequence: Because the value of a piece part used to fix an SRA is expressed through E(AWP) for the SRA, piece parts used to fix SRAs will be undervalued. If the approach described here were extended to lower levels of indenture, each lower level of indenture would be penalized successively more because each level of indenture is aggregated by the tall-pole formula. Unfortunately, the paper tests described below did not test this aspect of the tall-pole formula.

It should be possible to alter the tall-pole formula to reduce this problem, although we have not done it. One approach would be to use a fudge factor Ψ such that $E(AWP) \approx \Psi T$, where T is the approximate expected AWP from the tall-pole formula. The fudge factor Ψ could be obtained by using simulations to get "exact" values of E(AWP) and then doing a linear regression of E(AWP)/T on things like the number of items, the variance of their waiting times, and their RFs.

How much does the tall-pole formula affect evaluation of investments? We examined this issue for a simplified version of the long-run problem in which WRAs have only piece parts on them. (Thus, the paper tests to follow do not provide information about how much the tall-pole formula understates the value of piece parts used to fix SRAs.) We did two separate paper tests.

First, we computed "exact" optimal sequences of stockage decisions—i.e., decisions at each iteration of the stockage algorithm—using Monte Carlo simulation, because genuinely exact calculations are impractical.[1] For a set of WRAs to be described below, we computed the efficiency of the decisions made by the tall-pole formula relative to the decisions made using "exact" expected AWP times.[2] The set of WRAs was created from two years of data from the Naval Industrial Materiel Management System (NIMMS). We used 73 WRAs on the AWG-9 radar and:

---

[1]This work was done by Jeffrey Stein.

[2]Efficiency is the saving in expected AWP time caused by the stockage decision made by the tall-pole computation, divided by the saving in expected AWP time caused by the best selection the greedy algorithm could have made, according to the simulation.

* Used as RFs the number of demands divided by the number of jobs;
* Used as waiting times the average of the waits in the NIMMS file;[3]
* Used unit prices from the NIMMS records.

The results were encouraging. For most of these 73 WRAs, the tall-pole formula's efficiency exceeded 0.95 for the entire sequence of stockage decisions. For a few WRAs, the tall-pole's efficiency was lower for the early stockage decisions but still exceeded 0.8. The tall-pole formula was inefficient for four WRAs, all of which had two or more parts with RF $p_i = 1$ and sharing the maximum W_i. In such situations, the tall-pole formula gives zero return for all parts and does not stock anything. Several heuristics avoid this problem, but we cannot recommend any as clearly the best. Here are two:

* Stock one unit of each of the parts having $p_i = 1$ and the largest W_i; or
* Reduce the p_i to something like 0.999.

The second paper test of the effect of the tall-pole formula is described in MR-313-A/USN, *An Approach to Understanding the Relative Value of Parts*, by M.K. Brauner, J.S. Hodges, and D.A. Relles, forthcoming, and will be summarized here. We applied the long-run method to all WRAs in the 11 SMICs with the most activity in a NIMMS data set covering three years at three NADEPs. We also used the NIMMS data set to construct bills-of-material and to estimate WRA induction rates, expected OSTs, and RFs. The estimation methods we used are discussed in Brauner, Hodges, and Relles, op. cit.

We did the following assessment for each SMIC. We began with zero authorized stock of repair parts for the WRAs in the SMIC, then we constructed ranked lists of several thousand increments to authorized stock, using the U assumptions. We defined "baseline" authorized stock by cutting off the ranked list so that the resulting expected AWP time approximately matched the AWP time observed in the data. (We computed expected AWP time using the non-parametric simulation to be described below.) We then defined "treatment" authorized stock levels to be the baseline stock plus the next 1000 units on the ranked list of increments to authorized stock. We then computed the rate of return for treatment over baseline using

* The return computed by our methods (using the tall-pole and the U assumptions),
* The return estimated from a parametric simulation (described below), and
* The return estimated from a non-parametric simulation (described below).

Both the parametric and non-parametric simulations represented induction of WRA jobs, requisitions for parts, servicing of requisitions from depot stocks or from the supply system, and replenishment of depot stocks. The parametric simulation did so using a stochastic model identical to the U assumptions except without SRA repair. The parameters needed to run this simulation were estimated from the NIMMS data. The non-parametric simulation used the NIMMS data more directly: it used

---

[3]These waiting times do not describe the behavior of the supply system because they are affected by the stockage policy used by the depots from which the data were obtained. However, for the present purpose, they are adequate.

- The induction date (actually, the date of the first requisition) of each WRA job,
- The parts requisitioned, and
- The time the supply system took to service requisitions.[4]

The parametric and non-parametric simulations test different aspects of our methods. Both used the same definition of AWP time as our methods, so that key definition was not tested. The parametric simulation followed the U assumptions, so it was a pure test of the effect of the tall-pole formula. The non-parametric simulation was more demanding in that it did not follow the U assumptions regarding induction of WRA jobs and OST.

Again, the results are encouraging. For the 11 SMICs, the three computations produced similar rates of return aggregated to the SMIC level and disaggregated to the WRA level. Thus, although the tall-pole formula understates the AWP time, it does not seem to bias the computation of the *change* in AWP time arising from changes in authorized stock.

B.2. ASSUMPTIONS FOR THE SHORT-RUN PROBLEM

B.2.1. List of Assumptions

The following list categorizes the assumptions for the short-run method according to their function.

Definitions:

- Each unit on the WRA repair schedule makes a known contribution to aircraft availability at a specified time; the contributions of different units are additive.
- To define AWP time:
  —all part needs are known at induction
  —all needed parts are needed at induction.
- WRA and SRA turnaround time = AWP + process time, with part installation time included in process time.
- SRA repairs require only piece parts.

For ease of computation:

- The approximation in Expression (5.3.4).

Data availability:

- Accurate bills-of-material, including RFs for piece parts and SRAs.
- A menu of available supply actions.
- Costs of supply actions.
- The current supply position: for each part, the number on-shelf and the due-in schedule.
- For each WRA, on-aircraft installation time, forward shipping time, and process time.

[4]Actually, only about 40 percent of requisitions in the NIMMS data were serviced directly from the supply system and had real service times. The other requisitions were serviced from NIF stores, and we imputed system service times for them by making a draw from a probability distribution estimated from requisitions that *were* serviced by the supply system.

To facilitate inventory theory:

- The depot has adequate repair capacity (infinite servers).
- The induction rate of a WRA is independent of the number of units in the repair pipeline.
- No condemnations of SRAs or WRAs.
- No common items.
- No cannibalization.
- FIFO (service in order of arrival).
- WRA inductions and OST are stochastically independent.
- UPA = 1.
- (S – s, S) ordering.
- Induction times of WRA jobs are specified.
- OST from the supply system is constant and deterministic.
- Demands for parts to fix WRAs and SRAs are Bernoulli; the probability a part will be demanded in a job is its RF; parts are independent of one another.

B.2.2. Discussion of the Effects of the Assumptions

We have not paper-tested the short-run method, so we know little about the effects of the assumptions. As with the long-run method, we believe that the assumptions used to define AWP time are important, and their effects are unknown. Data requirements are greater for the short-run methods, but as with the long-run methods, we do not know the effects of degraded data. The comments made in Section B.1.2 regarding the miscellaneous assumptions should hold for the short-run problem as well.